UNIVERSITY OF AMSTERDAM

UvA-DARE (Digital Academic Repository)

Fair or Unfair Differentiation?

Reconsidering the Concept of Equality for the Regulation of Algorithmically Guided Decision-Making

Naudts, L.P.A.

Publication date
2023

Link to publication

Citation for published version (APA):
Naudts, L. P. A. (2023). *Fair or Unfair Differentiation? Reconsidering the Concept of Equality for the Regulation of Algorithmically Guided Decision-Making*. [Thesis, fully external, KU Leuven].

General rights
It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

Disclaimer/Complaints regulations
If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: https://uba.uva.nl/en/contact, or a letter to: Library of the University of Amsterdam, Secretariat, P.O. Box 19185, 1000 GD Amsterdam, The Netherlands. You will be contacted as soon as possible.

UvA-DARE is a service provided by the library of the University of Amsterdam (https://dare.uva.nl)

Download date:06 Oct 2026

KU Leuven
Faculty of Law and
Criminology

FAIR OR UNFAIR DIFFERENTIATION?

RECONSIDERING THE CONCEPT OF EQUALITY FOR THE REGULATION OF ALGORITHMICALLY GUIDED DECISION-MAKING

Laurens NAUDTS

Supervisor:
Prof. A. Vedder
Co-supervisor:
Prof. S. Sottiaux

Dissertation presented in
partial fulfillment of the
requirements for the degree of
Doctor in Law

January 2023

KU Leuven
Faculteit Rechtsgeleerdheid en
Criminologische Wetenschappen

EERLIJKE OF ONEERLIJKE DIFFERENTIATIE?

EEN HEROVERWEGING VAN HET BEGRIP GELIJKHEID TER REGULERING VAN ALGORITMISCH GEÏNFORMEERDE BESLUITVORMING

Laurens NAUDTS

Promotor:
Prof. A. Vedder
Co-Promotor:
Prof. S. Sottiaux

Proefschrift voorgedragen
tot het behalen van de
graad van Doctor in de
Rechten

Januari 2023

Table of Contents

MEMBERS OF THE EXAMINATION COMMITTEE I

ACKNOWLEDGEMENTS II

ABSTRACT V

INTRODUCTION: RECONSIDERING THE CONCEPT OF EQUALITY FOR THE REGULATION OF ALGORITMICALLY GUIDED DECISION-MAKING? I

RESEARCH CONTEXT, QUESTION AND METHODOLOGY II

A. Research Context ii

B. Research Question, Objectives and Contribution viii

C. Research Methodology, Structure and Outline xi

Part I - (In)equality in the Algorithmic Environment: Identifying The Risk of Egalitarian Harm. xi

Part II - Evaluating the Law: The Legal Conceptualisation of Equality and the Risk of Algorithmic Egalitarian Harm. xv

Part III - Navigating Political Philosophy: Socio-Relational and Distributive Conceptualisations of Equality and the Articulation of Algorithmic Egalitarian Harm. xviii

Part IV – Reconsidering the Concept of Equality for the Evaluation and Regulation of Algorithmically Guided Decision-Making: Concluding Remarks and Recommendations. xx

D. A Word on the Notion of Fairness xxi

PART I – (IN)EQUALITY IN THE ALGORITHMIC ENVIRONMENT: IDENTIFYING THE RISK OF EGALITARIAN HARM. 1

1 ALGORITHMICALLY GUIDED DECISION-MAKING: DIFFERENTIATION AND INEQUALITY 2

1.1 The Functions of Algorithms 2

1.1.1 Algorithms 3

1.1.2 Algorithmic Differentiation and (In)equality 8

1.2 Why Treat Algorithmically Guided Decisions Differently? 12

1.2.1 Differences in Reasoning 13

1.2.2 The Complexity of Decision-Making Grounds 15
1.2.3 Group-level Knowledge 16
1.2.4 Scale and Amplification 19
1.2.5 The Black Box: An Obstacle Towards Evaluation? 21

1.3 Algorithmic Power: Infrastructures, Decisional Control and Knowledge Creation. 23
1.3.1 Infrastructural Control 24
1.3.2 Decisional Control 26
1.3.3 Control over Knowledge 27

1.4 Conclusion 28

2 EQUALITY AS A MULTIDIMENSIONAL CONCEPT AND THE RISK OF ALGORITHMIC EGALITARIAN HARM 30

2.1 Equality: A Multidimensional Concept 31
2.1.1 The Moral Dimension 33
2.1.2 The Socio-Relational Dimension 36
2.1.3 The Distributive Dimension 38
2.1.4 Dimensional Interdependence 40

2.2 Algorithmic Inequality: The Risk of Egalitarian Harm 43
2.2.1 Algorithmic Egalitarian Harms 43
Historical Disadvantage 44
Algorithmic Moral Harms 47
Algorithmic Socio-Relational Harms 47
Algorithmic Distributive Harms 52
2.2.2 Algorithmic Equality of Whom? 55

2.3 Algorithmic Inequality: Related Notions 58
Bias and Inequality 58
Equality and (Non-)Discrimination 60
Equality and Intersectional Differentiation 62

2.4 The Evaluative Advantage of Equality 66

2.5 Conclusion 70

PART II – EVALUATING THE LAW: THE LEGAL CONCEPTUALISATION OF EQUALITY AND THE RISK OF ALGORITHMIC EGALITARIAN HARM. 73

INTRODUCTION **74**

3 UNCOVERING THE LEGAL CONCEPTUALISATION OF EQUALITY **76**

3.1 Equality: A Multi-Layered Legal Notion **76**
3.1.1 Models of Equality 77
3.1.2 Formal and Substantive Approaches Towards Equality 79

3.2 The Legal Conceptualisation of Equality: Two Lines of Inquiry **82**
3.2.1 The (Initial) Range of Protection: Open or Closed Models of Equality? 83
3.2.2 The Justification of Differentiation 85
3.2.3 Grounds or Principles? 86

3.3 Conclusion **89**

4 EQUALITY AND NON-DISCRIMINATION UNDER THE EUROPEAN CONVENTION ON HUMAN RIGHTS **91**

4.1 The Legal Framework **91**

4.2 The (Initial) Range of Protection **94**
An Open, Grounds-Based Approach Made Conditional? 94

4.3 The Justification of Differentiation **102**
4.3.1 The Review Procedure 103
Comparability 103
Legitimate Aim 107
Proportionality 110
4.3.2 Normative Guidance? 114
European Consensus and Common Ground 115
The 'Better Placed' Argument, or the Domain of Application 117
Democracy, dignity, and autonomy 120
Immutability 124
Irrelevance 128
The Anti-stereotyping Approach 129
Vulnerability 134
Proportionality as a Fail-Safe? 139

4.4 Conclusion **139**

5 EQUALITY AND NON-DISCRIMINATION IN THE EUROPEAN UNION **143**

5.1 The Legal Framework 143
5.1.1 The EU Equality Directives 146
5.1.2 The EU Charter of Fundamental Rights 147

5.2 The (Initial) Range of Protection 154
5.2.1 The EU Equality Directives 154
Extending the Meaning of Differentiation Grounds 156
Extending the Number of Differentiation Grounds through the Charter 158
5.2.2 The EU Charter of Fundamental Rights 162
Article 20 CFR 162
Article 21 CFR 164

5.3 The Justification of Differentiation 167
5.3.1 Justification in the EU Equality Directives 168
The Justification of Direct Discrimination 171
The Justification of Indirect Discrimination 173
5.3.2 The EU's Equality Review 174
The Standard Formula 175
Open-ended Questions Regarding the Application of the Charter 184
5.3.3 Normative Guidance? 185
Dignity and Autonomy 188
Irrelevance and Immutability 189
Anti-Stereotyping and Vulnerability 191
Rationality, Arbitrariness and Reliability 193
Articulation and Proportionality 201

5.4 Conclusion 202

PART III – NAVIGATING POLITICAL PHILOSOPHY: SOCIO-RELATIONAL AND DISTRIBUTIVE CONCEPTUALISATIONS OF EQUALITY AND THE ARTICULATION OF ALGORITHMIC EGALITARIAN HARM. 206

INTRODUCTION 207

6 SOCIO-RELATIONAL CONCEPTUALISATIONS OF EQUALITY AND ALGORITHMICALLY GUIDED DECISION-MAKING 209

6.1 Algorithmic Oppression and Domination 211
6.1.1 Oppression and Domination: Vertical and Horizontal Phenomena 213
6.1.2 A Function of Power? 214

6.1.3 Socio-Relational Status Harms: A Type of Vulnerability 216
6.1.4 Equality as Freedom from Algorithmic Domination and Oppression? 218
6.1.5 Algorithmic Oppression 219
Exploitation of the Powerless 220
Algorithmic Marginalisation: Data Exclusion and Over-Exposure 223
Algorithms and Cultural Imperialism 226
Algorithms and Violence 229
6.1.6 Algorithmic Domination 230
What Choice? 232
Interference in (the resources) of Choice 236
Active and Passive Domination 237

6.2 Algorithmic Stereotyping and Generalisations 243
6.2.1 Could (Novel) Algorithmic Stereotypes Arise? 245
6.2.2 The Problematic Nature of Generalisations: Deindividualisation and Misrepresentation. 247

6.3 Conclusion 251

7 DISTRIBUTIVE CONCEPTUALISATIONS OF EQUALITY AND ALGORITHMICALLY GUIDED DECISION-MAKING 254

7.1 Introduction 254

7.2 Equality of Primary Goods 256
7.2.1 People as Free and Equal: Two Moral Powers 257
7.2.2 The Principles of Justice 260
Equal Basic Liberties 261
Fair Equality of Opportunity 264
The Difference Principle 269
(The Social Bases of) Self-Respect 272

7.3 Equality as Capability 275
7.3.1 Capabilities and Functionings 278
7.3.2 Equal Human Dignity: Equality of Capability 280
7.3.3 The Central Human Capabilities 283
7.3.4 The Central Capabilities: An Evaluative Benchmark for Algorithmically Guided Decision-Making 287
7.3.5 The Central Capabilities: A Transformative Approach towards Algorithmic Decision-Making? 293

7.4 Conclusion 298

PART IV- RECONSIDERING THE CONCEPT OF EQUALITY FOR THE EVALUATION AND REGULATION OF ALGORITHMICALLY GUIDED DECISION-MAKING: CONCLUDING REMARKS AND RECOMMENDATIONS 302

8 REPOSITIONING THE CONCEPT OF EQUALITY FOR THE REGULATION OF ALGORITHMICALLY GUIDED DECISION-MAKING: CONCLUDING REMARKS AND RECOMMENDATIONS 303

8.1 Research Context, Question and Methodology 303

8.2 Main Research Findings 305

8.2.1 Part I - (In)equality in the Algorithmic Environment: Identifying The Risk of Egalitarian Harm. 305

The Algorithmic Environment 305

Equality as a Multidimensional Concept 307

Identifying Egalitarian Harm 308

8.2.2 Part II - Evaluating the Law: The Legal Conceptualisation of Equality and the Risk of Algorithmic Egalitarian Harm. 310

8.2.3 Part III - Navigating Political Philosophy: Socio-Relational and Distributive Conceptualisations of Equality and the Articulation of Algorithmic Egalitarian Harm. 316

8.2.4 Part IV - Reconsidering the Concept of Equality for the Evaluation and Regulation of Algorithmically Guided Decision-Making. 322

Equality as Defined through a Balancing of Interests? 323

Motivating the Balancing Review 325

The Importance of Articulation 328

Fundamental Interests as the Point of Departure 329

Equality: A Notion in Need of Support 331

BIBLIOGRAPHY 334

Laws 334

Case-Law 335

European Court of Human Rights 335

European Court of Justice 337

Journals, Books, Reports and Miscellaneous 338

Members of the Examination Committee

Prof. Dr. Anton Vedder (Supervisor)

Prof. Dr. Stefan Sottiaux (Co-supervisor)

Prof. Dr. Elise Muir

Prof. Dr. Wibren van der Burg

Emer. Prof. Dr. Yves Poullet

Emer. Prof. Dr. Jacques Herbots (Chairman)

Acknowledgements

The final year of my doctoral research was one of hardship and reflection. To write these words then on the precipice of the new year, is an experience of relief and hope. With them I can turn the page and look forward to what lies ahead with optimism. Though this day marks the end of a long, arduous, and anxiety-ridden journey, both as a researcher and a human being I remain a work in progress. That I was able to arrive at my destination was only made possible through the support and encouragement of many, whom I would like to thank.

To my supervisor, Prof. Dr. Anton Vedder. While I might not have mastered the art of succinctness quite as much as you would have liked me to, do know that I tried! I have multiple documents on my computer entitled “leftovers”. Having been told to be concise on multiple occassions, I understand that a simple “thank you” would suffice. Please allow me to nonetheless add a few more words as a token of my appreciation. When I sent you my first rough and perhaps slightly futuristic research proposal on big data, simulations and policy making, you immediately saw its potential. You have given me the prerogative to explore, study and wonder. You were strict when you had to, kind when you needed to. Our many conversations, and sometimes confrontations, have made me a better, well-rounded researcher.

To my co-supervisor Prof. Dr. Stefan Sottiaux. You always pushed me to greater clarity. You identified, and made me reconsider, the preconceived notions I had concerning the problems I was trying to solve and the role and functioning of the law for doing so. After our discussions, I often found myself lost in thought. Those moments of introspection have greatly contributed to my formation as a scholar.

I would like to thank the members of my supervisory and exam committee for their observations, feedback, and valid criticisms. To Prof. Dr. Wibren van der Burg, Emer. Prof. Dr. Yves Poullet and Prof. Dr. Elise Muir. Your knowledge and expertise have sharpened my mind. You took the time to engage with my research with great interest and enthusiasm. I would like to thank Emer. Prof. Dr. Jacques Herbots for acting as the chairman of my doctoral jury and defence. Just moments after my predefence, we met for the first time. Upon my introduction, you jovially tapped my shoulder and exclaimed: “Congratulations, you did it!” Though this encounter was rather brief, it was one of great exaltation.

To Prof. Dr. Bettina Berendt. I am indebted to your willingness to engage with me in endless conversations regarding the social impact of algorithmic decision-making systems, unafraid to

correct my technological assumptions in the process. Our conversations on equality and injustice were insightful and motivational. I am also thankful for the opportunity you gave me to join your research group at the Berlin Weizenbaum institute as a research fellow. During my stay, you introduced me to an equally engaged group of young researchers. To Doctors Milagros Miceli and Diana Alina Serbanescu. Your research, interdisciplinarity and openness have been inspirational. The talks we shared on power in the algorithmic environment have been formative for my research and I am grateful that I was able to get to know, and collaborate with, you.

To Prof. Dr. Peggy Valcke. When I first arrived at CiTiP, you showed me the ropes. I fondly remember our first work trip and deliverable – you even managed to sneak in a Yogi Berra quote: "It's tough to make predictions, especially about the future"! It was the start of a great collegial relationship. It was an honour to coordinate and help develop the IusStart incubator programme under your tutelage. I hope it can – and a prediction I am willing to make - continue to inspire new generations of students.

To my colleagues at CiTiP. Over the past years, I have had the privilege and opportunity to meet and get to know a great many of you. For these relationships I am grateful. I have learned a lot from you. I treasure your knowledge, solidarity and inclusivity: you are the beating heart of our centre. To Shuki. You are one-of-a-kind. CiTiP has heart, but you are the electric pulse that keeps the whole thing going. You are gentle and supportive. In the often-high-pressure environment of academia, you were a piece of tranquillity whose ear I could always bend.

To my friends. Thank you for sticking with me: I promise to be more social again.

To my dad and my brother Mathias. Your concern for my happiness has been a continuous source of comfort and guidance. In times I was lost, you have given me direction. You take the time to listen to my grievances, concerns and doubts. We can discuss, argue and laugh, and in moments of sadness, we can find love and joy. I could not wish for a better family.

To Plixavra. You are without doubt the wisest and kindest person I know. Whether it be law, politics, art, movies, or emotions, there is no conversation or worry we cannot share. You are my lifeline in times of anxiety. When I am too much (I am well aware), you always remain understanding, compassionate and patient. I could spend another four years and threehundred plus pages putting into words how much you mean to me. Though I imagine you wouldn't want to read another book of mine anytime soon, for all that you have done for me, you would be entirely

deserving. You are a gift to my life and I hope I can be the same to you. Till then my gratitude is endless.

To my mother. Moeke, your dedication to my well-being and happiness has been unfailing. Every step of the way, you were there for me and I would not be standing here today without you. I will always cherish our time together and your memory. I miss you tremendously. I love you. I dedicate this book to you.

Laurens Naudts
December 2022

Abstract

Algorithms are increasingly relied upon to help decision-makers automate, streamline, structure and guide a variety of decision-making processes, ranging from trivial to critical, in both the public and private sector. This dissertation is concerned with a specific type of injustice that may come along with the development and deployment of algorithmically guided decision-making systems: the emergence of unjust (in)equality brought about by various instances of differentiation and differential treatment that take place within and as part of these systems.

In this data-driven environment, people and groups of people are continuously classified, categorised, ranked and scored on a variety of features or attributes, such as their characteristics, interests, behaviour and preferences. For decision-subjects, these differentiation acts can generate significant consequences regarding their social position and life's prospects: they affect the choices and options they are presented, the interactions and relationships they hold with others and themselves, the opportunities they are given, the burdens and benefits they carry, and so forth. Yet, when applied on a large enough scale, these decisions may also initiate significant societal change.

Due to the complexity of the digital environment and the distinctive characteristics algorithmically guided decisions exhibit, it has become increasingly difficult to assess whether the decisions these knowledge and data-driven systems inform, and the (in)equalities they produce, can be justified. As a consequence, uncertainty exists as to whether current and future regulatory efforts, and the approach to equality they (will) adhere to, have the capacity to respond to the egalitarian harms algorithms risk to introduce in an appropriate manner.

In this dissertation, I reposition and operationalise the notion of equality as a practicable and interpretative lens to strengthen the evaluation and regulation of algorithmically guided decision-making practices in light of the inequalities they (risk to) produce.

This dissertation begins with a definition of the algorithmic research context in which the notion of equality will be operationalised. I explore a series of characteristics that typify algorithmic systems and render the inequalities they generate distinctive in terms of their form and scope. Due to these unique characteristics, algorithmic inequalities have the potential to restructure the fabric of society alongside new and existing dimensions: they may not only reinforce existing social injustice, but they may also introduce new forms of non-representational injustice **(Chapter 1).**

Drawing inspiration from both European equality and non-discrimination law and political philosophical theories of justice, and informed by the (practical) functioning of algorithmic decision-making systems and the particular challenges these bring along, I propose equality as a multidimensional concept that can be specified alongside three (interrelated) axes. The model represents a core set of ideals commonly associated with the notion of equality as a social value: equal concern and respect (the moral dimension), equal social standing and equal social relationships (the socio-relational dimension), and/or equal access to certain justice-relevant goods (the distributive dimension) **(Chapter 2)**. Throughout this dissertation, this multidimensional understanding of equality is operationalised to identify, articulate and evaluate algorithmic injustice, and the response formulated thereto within a given policy, law, code or theory.

In a first step, my understanding of equality is positioned against the algorithmic environment in order to articulate and identify the egalitarian harms algorithms risk to impose onto decision-subjects and society at large **(Part I: Identification, Chapter 2)**.

In a second step, the multidimensional model functions as a support mechanism to uncover and evaluate the legal conceptualisation of equality found within European equality and non-discrimination law (Council of Europe and European Union). By investigating whom the law protects (equality of whom?) against which inequalities (what egalitarian ideals do they promote?) and how (what review mechanisms have been put in place when equality is positioned against competing values?), I assess whether the legal approach to equality can address the injustice algorithms risk to introduce **(Part II: Evaluation; Chapters 3-5).**

In a third step, I rely upon the model to locate and examine specific notions of equality – chosen due to their correspondence with the aforementioned dimensions. These socio-relational (domination and oppression) and distributive (primary goods and capabilities) notions are examined to concretise the egalitarian harms algorithms risk to produce, the conditions under which these harms may manifest in practice, and the safeguards that can be provided to protect decision-subjects and society at large against them **(Part III: Navigation; Chapter 6 and 7).**

Finally, based upon my findings, I formulate a set of normative recommendations aimed at repositioning the concept of equality within the algorithmic governance debate in an effort to strengthen its guiding function for the evaluation and regulation of algorithmically informed decision-making systems **(Part IV: Synthesis; Chapter 8)**.

Introduction: Reconsidering the Concept of Equality for the Regulation of Algoritmically Guided Decision-Making?

Research Context, Question and Methodology

A. Research Context

Algorithmically guided decision-making systems, and the actors who deploy them, have gained significant power in the arrangement of modern-day society. Algorithms have become ubiquitous. They automate, streamline, structure and guide a variety of decision-making processes – trivial and critical - in both the public and private sector. Algorithms are used to organise welfare allocation[1] or aid job seekers in finding their way back onto the employment market.[2] They can help identify when individuals are at risk, or conversely, when they pose a risk, in fields of

[1] *See* among others*:* Virginia Eubanks, *Automating Inequality: How High-Tech Tools Profile, Police and Punish the Poor* (Sint Martin's Press: Macmillan Publishers, 2018); Charlotte Huisman, 'SyRI, het fraudesysteem van de overheid, faalt: nog niet één fraudegeval opgespoord', de Volkskrant, 27 June 2019, https://www.volkskrant.nl/gs-b789bc3a; Koen Vervloesem, 'How Dutch Activists Got an Invasive Fraud Detection Algorithm Banned', AlgorithmWatch, 2020, https://algorithmwatch.org/en/story/syri-netherlands-algorithm/; Marvin van Bekkum and Frederik Zuiderveen Borgesius, 'Digital Welfare Fraud Detection and the Dutch SyRI Judgment', *European Journal of Social Security*, 2 August 2021, 13882627211031256, https://doi.org/10.1177/13882627211031257; Adamantia Rachovitsa and Niclas Johann, 'The Human Rights Implications of the Use of AI in the Digital Welfare State: Lessons Learned from the Dutch SyRI Case', *Human Rights Law Review* 22, no. 2 (1 June 2022): ngac010, https://doi.org/10.1093/hrlr/ngac010.

[2] See among others: Michael Veale, Max Van Kleek, and Reuben Binns, 'Fairness and Accountability Design Needs for Algorithmic Support in High-Stakes Public Sector Decision-Making', *ArXiv:1802.01029 [Cs]*, 2018, 1–14, https://doi.org/10.1145/3173574.3174014; Nicolas Kayser-Bril, 'Austria's Employment Agency Rolls out Discriminatory Algorithm, Sees No Problem', *AlgorithmWatch* (blog), 6 October 2019, https://algorithmwatch.org/en/story/austrias-employment-agency-ams-rolls-out-discriminatory-algorithm/; Paola Lopez, 'Reinforcing Intersectional Inequality via the AMS Algorithm in Austria', in *Proceedings of the STS Conference Graz* (STS Conference Graz, Graz, 2019), 21; Paola Lopez, 'Bias Does Not Equal Bias: A Socio-Technical Typology of Bias in Data-Based Algorithmic Systems', *Internet Policy Review* 10, no. 4 (7 December 2021), https://policyreview.info/articles/analysis/bias-does-not-equal-bias-socio-technical-typology-bias-data-based-algorithmic; Doris Allhutter et al., 'Algorithmic Profiling of Job Seekers in Austria: How Austerity Politics Are Made Effective', *Frontiers in Big Data* 3 (2020): 5, https://doi.org/10.3389/fdata.2020.00005.

medicine[3], law enforcement and intelligence, criminal sentencing[4], and beyond. They can guide hiring decisions[5], credit and insurance applications[6] or education allocation.[7] They recommend what movies to watch, music to listen to, who to follow and interact with on social media and what

---

[3] See among others*:* Anamika Barman-Adhikari et al., 'Social Networking Technology Use and Engagement in HIV Related Risk and Protective Behaviors among Homeless Youth', *Journal of Health Communication* 21, no. 7 (July 2016): 809–17, https://doi.org/10.1080/10810730.2016.1177139; Zi-Yi Dou et al., 'Harnessing Social Media to Identify Homeless Youth At-Risk of Substance Use', *Proceedings of the AAAI Conference on Artificial Intelligence* 35, no. 17 (18 May 2021): 14748–56.; Shannon L. Stewart et al., 'Risk of Injury to Others: The Development of an Algorithm to Identify Children and Youth at High-Risk of Aggressive Behaviours', *Frontiers in Psychiatry* 12 (3 January 2022): 750625, https://doi.org/10.3389/fpsyt.2021.750625.

[4] See among others: ProPublica Data Store, 'COMPAS Recidivism Risk Score Data and Analysis', text/html, ProPublica Data Store (ProPublica Data Store, 2 May 2016), https://www.propublica.org/datastore/, https://www.propublica.org/datastore/dataset/compas-recidivism-risk-score-data-and-analysis; Surya Mattu et al., 'How We Analyzed the COMPAS Recidivism Algorithm', ProPublica, 23 May 2016, https://www.propublica.org/article/how-we-analyzed-the-compas-recidivism-algorithm?token=e3k6mE9frXgAOt3zvKfhyLDrUGgrhmOb; Danielle Ensign et al., 'Runaway Feedback Loops in Predictive Policing', *ArXiv:1706.09847 [Cs, Stat]*, 29 June 2017, http://arxiv.org/abs/1706.09847; Plixavra Vogiatzoglou, 'Mass Surveillance, Predictive Policing and the Implementation of the CJEU and ECtHR Requirement of Objectivity', *European Journal of Law and Technology* 10, no. 1 (16 May 2019), https://ejlt.org/index.php/ejlt/article/view/669; Laurens Naudts, 'Criminal Profiling and Non-Discrimination: On Firm Grounds for the Digital Era?', in *Security and Law. Legal and Ethical Aspects of Public Security, Cyber Security and Critical Infrastructure Security.*, ed. Anton Vedder et al. (Camdridge, Antwerp, Chicago: Intersentia, 2019), 63–96.

[5] See for example: Maarten Buyl et al., 'Tackling Algorithmic Disability Discrimination in the Hiring Process: An Ethical, Legal and Technical Analysis', in *2022 ACM Conference on Fairness, Accountability, and Transparency* (FAccT '22: 2022 ACM Conference on Fairness, Accountability, and Transparency, Seoul Republic of Korea: ACM, 2022), 1071–82, https://doi.org/10.1145/3531146.3533169; Gideon Mann and Cathy O'Neil, 'Hiring Algorithms Are Not Neutral', *Harvard Business Review*, 9 December 2016, https://hbr.org/2016/12/hiring-algorithms-are-not-neutral; Aislinn Kelly-Lyth, 'Challenging Biased Hiring Algorithms', *Oxford Journal of Legal Studies* 41, no. 4 (1 December 2021): 899–928, https://doi.org/10.1093/ojls/gqab006.

[6] See for instance: Mikella Hurley and Julius Adebayo, 'Credit Scoring in the Era of Big Data', *Yale J.L. & Tech* 18 (2016): 148–216; Teresa Bono, Karen Croxson, and Adam Giles, 'Algorithmic Fairness in Credit Scoring', *Oxford Review of Economic Policy* 37, no. 3 (23 September 2021): 585–617, https://doi.org/10.1093/oxrep/grab020.

[7] Alex Hern, 'Ofqual's A-Level Algorithm: Why Did It Fail to Make the Grade?', the Guardian, 21 August 2020, http://www.theguardian.com/education/2020/aug/21/ofqual-exams-algorithm-why-did-it-fail-make-grade-a-levels; Heather Stewart, Sally Weale, and Kate Proctor, 'Ofqual Chief to Face MPs over Exams Fiasco and Botched Algorithm Grading', *The Guardian*, 20 August 2020, sec. Education, https://www.theguardian.com/education/2020/aug/20/ofqual-chief-to-face-mps-over-exams-fiasco-and-botched-algorithm-grading-system; International Baccalaureate, 'IB Assessment for May 2020 Examination Session during COVID-19 Pandemic', 2020, https://www.ibo.org/globalassets/news-assets/coronavirus/assessment-model-letter-may-2020-en.pdf; Hye Jung Han, 'An Algorithm Shouldn't Decide a Student's Future', POLITICO, 13 August 2020, https://www.politico.eu/article/an-algorithm-shouldnt-decide-students-future-coronavirus-international-baccalaureate/.

political news people will read.[8] Algorithms help businesses in determining who to target during advertisement campaigns or what prices can be charged.[9] Algorithmically guided decision-making instances rely upon differentiation. People and groups of people are classified, categorised, ranked and scored, on a variety of features or attributes, such as their characteristics, interests, behaviour, and preferences: Who is the ideal candidate to hire? Who should we give a loan? Who should we connect online? What options should we display to whom? And so forth.[10]

This dissertation is concerned with a specific type of injustice that may come along with the development and deployment of algorithmically guided decision-making systems: the emergence of unjust (in)equality brought about by various instances of differentiation and differential treatment that take place within and as part of these systems. As illustrated by the examples below, for those subject to an algorithmically guided decision, classification acts can have a substantial impact as they affect (among others) the choices and options they are presented with, the interactions and relationships they hold with others and themselves, as well as the opportunities, burdens, benefits and risks they are afforded.[11]

In an effort to raise the efficiency and effectiveness of the Austrian public employment service (AMS), an algorithm was built that would automatically classify job seekers into one of three categories depending on their outlook of finding employment, ranging from high prospects in the short term to low prospects in the long term.[12] The system was criticised due to the presence of gender, disability and age (among other ascribed characteristics) as criteria during the allocation

---

[8] Shlomo Berkovsky, Ronnie Taib, and Dan Conway, 'How to Recommend ? User Trust Factors in Movie Recommender Systems', in *Proceedings of the 22nd International Conference on Intelligent User Interfaces IUI '17*, 2017, 287–300, https://doi.org/10.1145/3025171.3025209; Jaron Harambam, Natali Helberger, and Joris van Hoboken, 'Democratizing Algorithmic News Recommenders: How to Materialize Voice in a Technologically Saturated Media Ecosystem', *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences* 376, no. 2133 (28 November 2018): 20180088, https://doi.org/10.1098/rsta.2018.0088; Natali Helberger, Kari Karppinen, and Lucia D'Acunto, 'Exposure Diversity as a Design Principle for Recommender Systems', *Information, Communication & Society* 21, no. 2 (February 2018): 191–207, https://doi.org/10.1080/1369118X.2016.1271900; Ferenc Huszár et al., 'Algorithmic Amplification of Politics on Twitter', 2021, 27.

[9] See for instance: Frederik Zuiderveen Borgesius and Joost Poort, 'Online Price Discrimination and EU Data Privacy Law', *Journal of Consumer Policy* 40, no. 3 (1 September 2017): 347–66, https://doi.org/10.1007/s10603-017-9354-z; Tanya Kant, 'Identity, Advertising, and Algorithmic Targeting: Or How (Not) to Target Your "Ideal User"', *MIT Case Studies in Social and Ethical Responsibilities of Computing*, no. Summer 2021 (10 August 2021), https://doi.org/10.21428/2c646de5.929a7db6.

[10] In the context of this dissertation, the notions feature and attribute will be used as synonyms.

[11] These ideas will be further explored and expanded upon in Chapter 2 "Equality as an Evaluative Concept."

[12] Allhutter et al., 'Algorithmic Profiling of Job Seekers in Austria'.

process. Under one model, women, disabled people, and people over thirty would be weighed negatively by the system, and more likely classified as job seekers with no long-term prospects.[13] The latter group would receive less support from the AMS, and instead assigned to external agencies for improving their prospects.[14] This process was met with disapproval. The system was more likely to reinforce and perpetuate historical, socio-economic disadvantage and inequality on the job market.[15] Vulnerable and marginalised job seekers were deprived of, rather than provided the assistance they needed.

In 2020, amidst the COVID-19 pandemic, several educational institutions found themselves in the impossibility of organising final exam sessions. Alternative grading systems had to be introduced. The International Baccalaureate program was one of the organisations confronted with this challenge.[16] Because the final exams, which would typically constitute an essential element of a student's final grade, had to be cancelled, the International Baccalaureate developed a predictive algorithm. The algorithm would take into consideration the coursework scores students achieved, the expected grades that teachers predicted them to have and the school's historical results.[17] On the basis of these attributes the algorithm would predict the score students would likely have achieved if they had been in the possibility to take an exam.[18] The system was criticised for

[13] Kayser-Bril, 'Austria's Employment Agency Rolls out Discriminatory Algorithm, Sees No Problem'; Lopez, 'Reinforcing Intersectional Inequality via the AMS Algorithm in Austria'; Allhutter et al., 'Algorithmic Profiling of Job Seekers in Austria'. Moreover, the system would likely have the biggest impact for job seekers at the intersections of those attributes: a disabled women over the age of thirty. See on the latter point, also the analysis by Lopez in Lopez, 'Reinforcing Intersectional Inequality via the AMS Algorithm in Austria'. See also Chapter 2: Equality and Intersectional Discrimination.

[14] Allhutter et al., 'Algorithmic Profiling of Job Seekers in Austria', 7; Kayser-Bril, 'Austria's Employment Agency Rolls out Discriminatory Algorithm, Sees No Problem'. Lopez, 'Bias Does Not Equal Bias'.

[15] Allhutter et al., 'Algorithmic Profiling of Job Seekers in Austria', 7. Lopez, 'Bias Does Not Equal Bias'.

[16] Han, 'An Algorithm Shouldn't Decide a Student's Future'; 'The Norwegian DPA Intends to Order Rectification of IB Grades', Datatilsynet, 2020, https://www.datatilsynet.no/en/news/2020/the-norwegian-dpa-intends-to-order-rectification-of-ib-grades/; Datatilsynet (Norwegian Data Protection Authority), 'Advance Notification of Order to Rectify Unfairly Processed and Incorrect Personal Data - International Baccalaureate Organization', 7 August 2020, https://www.datatilsynet.no/contentassets/04df776f85f64562945f1d261b4add1b/advance-notification-of-order-to-rectify-unfairly-processed-and-incorrect-personal-data.pdf. See also: Hern, 'Ofqual's A-Level Algorithm'; Stewart, Weale, and Proctor, 'Ofqual Chief to Face MPs over Exams Fiasco and Botched Algorithm Grading'; BBC News, 'A-Levels and GCSEs: How Did the Exam Algorithm Work?', *BBC News*, 20 August 2020, sec. Explainers, https://www.bbc.com/news/explainers-53807730.

[17] Han, 'An Algorithm Shouldn't Decide a Student's Future'.

[18] Han.

insufficiently taking into account students' individual merit and potential as relevant features in determining fair equality of educational opportunity.

Full-body scanning algorithms have been deployed at airports to identify potential security threats. Once the body is scanned, the system will classify a person as posing a risk or no risk. When these systems are built with a non-inclusive and non-universal definition of the human body or gender in mind however, they are liable to misidentify transgender and gender nonconforming people as a potential security risk.[19] Similar warnings due to misidentification have also been triggered in relation to people's disability, religious headgear, or hairstyles.[20] In this particular case, the inequality imposed through classification is not socio-economic, but rather socio-relational in nature. A particular person or group of persons is misrecognised or misrepresented or, put differently, not recognised as an equal in the social relationships they enter. These acts can lower a person's self-respect and self-worth, and in turn, affect how they relate to themselves and towards others.

PRODUCERS OF (IN)EQUALITY - Algorithmic decision-making systems, and algorithmic classification acts, are producers of (in)equality. Consider for the moment a classic and often cited formulation of equality as a principle which mandates that like situations are treated alike and different situations are treated differently. (Algorithmic) decision-making processes share a strange relationship with this formulation of equality. First, algorithmic techniques are often deployed with the goal of detecting the similarities and differences between various situations within a given context. Second, in many decision-making contexts, this need to distinguish is seen as a welcome benefit. Decision-makers want to know who will enjoy what content, who qualifies for a welfare benefit, who is likely to commit an act of crime, etc. Imagine a hiring algorithm that will categorise applicants in two groups: ideal applicants and non-ideal applicants. Applicants with features a, b and c will be grouped as ideal, those with features -a, -b and -c will be placed in the opposite category. In other words, decision-making instances look to group together individuals in accordance with their similarities and in accordance with their differences. Members of the ideal group will be treated differently (or unequally) from members of the non-ideal group (more

[19] Lucas Waldron Medina Brenda, 'TSA's Body Scanners Are Gender Binary. Humans Are Not.', ProPublica, accessed 7 February 2022, https://www.propublica.org/article/tsa-transgender-travelers-scanners-invasive-searches-often-wait-on-the-other-side?token=7bjY-MRzWk5Ed4DCZRvFVYwt2HBrAFXd.

[20] The latter category was found to be particularly disadvantageous to women of colour. Medina.; Brenda Medina Frank Thomas, 'TSA Agents Say They're Not Discriminating Against Black Women, But Their Body Scanners Might Be', ProPublica, accessed 3 August 2022, https://www.propublica.org/article/tsa-not-discriminating-against-black-women-but-their-body-scanners-might-be.

likely to be hired versus not likely to be hired). At the same time, within each group, every member is subject to similar (or equal) treatment. Taken on its own, this formulation of equality appears normatively empty. It has not really told us much yet: When exactly are two situations alike? And should we guarantee like treatment every time a similarity is found between two situations?[21]

Furthermore, due to the increase of available data, the differentiation (or decisional) grounds used to generate these (in)equalities have become varied and diverse: they can be straightforward or complex, static or dynamic, tangible or intangible.[22] Given this expansion of differentiation grounds, and given the wide variety of domains in which algorithms can find application, legislators and regulators face difficulty when trying to assess whether data-driven knowledge systems can be rightfully used and applied: when, and under what conditions, are algorithmic decision-making systems fair? And to what extent should they intervene to mitigate unjustified differential treatment and the injustice this might entail for those involved? Yet, the more classification acts are performed within a given environment, the higher the risk that some of these forms of differentiation, and the inequalities that follow from them, are unjustified.[23]

---

[21] Peter Westen, 'The Empty Idea of Equality', *Harvard Law Review* 95, no. 3 (1982): 537–96; P. Westen, *Speaking of Equality: An Analysis of the Rhetorical Force of 'Equality' in Moral and Legal Discourse*, *Speaking of Equality* (Princeton University Press, 2014), https://doi.org/10.1515/9781400861484.

[22] See Section 1.2 Why Treat Algorithmically Guided Decisions Differently? See also: Nicholas Diakopoulos, 'Accountability in Algorithmic Decision Making', *Commun. ACM* 59, no. 2 (January 2016): 56–62, https://doi.org/10.1145/2844110. Brent Mittelstadt, 'From Individual to Group Privacy in Big Data Analytics', *Philosophy & Technology* 30, no. 4 (December 2017): 475–94, https://doi.org/10.1007/s13347-017-0253-7; Mireille Hildebrandt, 'Defining Profiling: A New Type of Knowledge?', in *Profiling the European Citizen: Cross-Disciplinary Perspectives*, ed. Mireille Hildebrandt and Serge Gutwirth (Dordrecht: Springer Netherlands, 2008), 17–45, https://doi.org/10.1007/978-1-4020-6914-7_2; Janneke Gerards and Frederik Zuiderveen Borgesius, 'Protected Grounds and the System of Non-Discrimination Law in the Context of Algorithmic Decision-Making and Artificial Intelligence', *Colorado Technology Law Journal* 20 (June 2022): 56; Sandra Wachter, 'The Theory of Artificial Immutability: Protecting Algorithmic Groups under Anti-Discrimination Law', *SSRN Electronic Journal*, 2022, https://doi.org/10.2139/ssrn.4099100. Anton Vedder and Laurens Naudts, 'Accountability for the Use of Algorithms in a Big Data Environment', *International Review of Law, Computers & Technology* 31, no. 2 (4 May 2017): 206–24, https://doi.org/10.1080/13600869.2017.1298547. Laurens Naudts, 'How Machine Learning Generates Unfair Inequalities and How Data Protection Instruments May Help in Mitigating Them', in *Data Protection and Privacy: The Internet of Bodies*, ed. Ronald Leenes et al. (Hart Publishing, 2018), 71–92; Laurens Naudts, 'AI En Algoritmische Categorisatie: Gelijkheid En Non-Discriminatie', in *Artificiële Intelligentie En Maatschappij*, ed. Jan De Bruyne and Nicolas Bouteca (Oud-Turnhout/'s-Hertogenbosch: Gompel & Svacina, 2021), 223–47.

[23] Of course, at the same time, the number of fair differentiation acts might go up as well. Vedder and Naudts, 'Accountability for the Use of Algorithms in a Big Data Environment', 218–19.

B. Research Question, Objectives and Contribution

As algorithms co-shape and mediate society, it is reasonable to question the use of these technologies and the conditions under which they are deployed. When algorithmically guided decisions interfere in people's lives, we want that this interference can be justified. In principle, the question whether the inequalities people face are justified, depends in large part upon the function and interpretation given to equality within the laws, policies or theories that govern the structure of society, and the way these frameworks believe the (competing) interests of people are best accommodated. A society's commitment towards equality reflects its intent to equalise the enjoyment of certain fundamental interests for all and societies can differ in how these interests are interpreted. Because algorithms co-govern social structures, they have the capacity to interfere with the operationalisation of equality. Due to their complexity however, it has become increasingly difficult to assess whether, and under what conditions, these data-driven systems, and the (in)equalities they generate, interfere with our environment in ways that are fair or unfair, justified, or unjustified. As a consequence, uncertainty exists as to whether current and future regulatory efforts, and the approach towards equality they (will) adhere to, have the capacity to respond to the egalitarian harms algorithms risk to introduce in an appropriate manner.[24]

---

[24] For example, the EU's High Level Expert Group on Artificial Intelligence's Guidelines on Trustworthy AI stipulate that individuals and groups should remain free from unfair bias and discrimination where AI applications are deployed. The Council of Europe's Ad Hoc Committee on Artificial Intelligence also recognised the need to ensure just algorithms, warning against the reinforcement of historical and current prejudice, including discrimination, as well as recognising the possibility for new forms of unfair differential treatment to emerge. In a report on automated decision-making commissioned by the European Parliament, Castelluccia and Le Metayer called upon the policy makers and scholars to articulate "beyond existing criteria already identified in anti-discrimination laws" other types of treatment that should also be considered undesirable?" High-Level Expert Group on Artificial Intelligence, 'Ethics Guidelines for Trustworthy AI', April 2019, 12, https://ec.europa.eu/digital-single-market/en/news/ethics-guidelines-trustworthy-ai. Ad Hoc Committee on Artificial Intelligence (CAHAI), 'Feasibility Study' (Council of Europe, 17 December 2020). Claude Castelluccia and Daniel Le Metayer, 'Understanding Algorithmic Decision-Making: Opportunities and Challenges' (European Parliament Research Service, March 2019), http://www.europarl.europa.eu/RegData/etudes/STUD/2019/624261/EPRS_STU(2019)624261_EN.pdf. For scholarship, see among others: Mireille Hildebrandt and Serge Gutwirth, eds., *Profiling the European Citizen: Cross-Disciplinary Perspectives* (New York: Springer, 2008); Anton Vedder, 'KDD: The Challenge to Individualism', *Ethics and Information Technology* 1, no. 4 (1999): 275–81; Bart Custers et al., eds., *Discrimination and Privacy in the Information Society: Data Mining and Profiling in Large Databases*, Studies in Applied Philosophy, Epistemology and Rational Ethics (Berlin Heidelberg: Springer-Verlag, 2013), https://www.springer.com/gp/book/9783642304866; Matthias Leese, 'The New Profiling: Algorithms, Black Boxes, and the Failure of Anti-Discriminatory Safeguards in the European Union', *Security Dialogue* 45, no. 5 (2014): 494–511, https://doi.org/10.1177/0967010614544204; Vedder and Naudts,

That being said, it is believed that a renewed investigation into the concept of equality can help clarify and revitalise the guiding function equality holds for the regulation of algorithmically guided decision-making, including the instances of differentiation that take place therein, and the inequalities that follow from them.

RESEARCH QUESTION – This dissertation's research question reads as follows:

How can we operationalise the concept of equality to guide the evaluation and regulation of algorithmically guided decision-making processes in light of the inequalities they produce?

The main research objective is therefore to reposition the notion of equality as a practicable and interpretative lens that can strengthen the evaluation and regulation of algorithmically guided decision-making practices in light of the inequalities they (risk to) produce. To that end, I develop a multidimensional understanding of equality, which will be operationalised through the performance of three particular functions, namely:

a) To identify and articulate when, and under what conditions, algorithmically guided decision-making processes, the instances of algorithmic differentiation that take place therein, and the inequalities that flow thereout, risk being fair or unfair, justified, or unjustified **(identification).**

'Accountability for the Use of Algorithms in a Big Data Environment'; Naudts, 'Criminal Profiling and Non-Discrimination: On Firm Grounds for the Digital Era?'; Naudts, 'How Machine Learning Generates Unfair Inequalities and How Data Protection Instruments May Help in Mitigating Them'; Monique Mann and Tobias Matzner, 'Challenging Algorithmic Profiling: The Limits of Data Protection and Anti-Discrimination in Responding to Emergent Discrimination', *Big Data & Society* 6, no. 2 (1 July 2019): 2053951719895805, https://doi.org/10.1177/2053951719895805; Sandra Wachter and Brent Mittelstadt, 'A Right to Reasonable Inferences: Re-Thinking Data Protection Law in the Age of Big Data and AI' 2019, no. 2 (2019): 494–620, https://doi.org/10.7916/d8-g10s-ka92; Sandra Wachter, 'Affinity Profiling and Discrimination by Association in Online Behavioral Advertising', *Berkeley Technology Law Journal* 35, no. 2 (2020): 367–430; Sandra Wachter, Brent Mittelstadt, and Chris Russell, 'Why Fairness Cannot Be Automated: Bridging the Gap Between EU Non-Discrimination Law and AI', SSRN Scholarly Paper (Rochester, NY: Social Science Research Network, 3 March 2020), https://doi.org/10.2139/ssrn.3547922; Frederik J. Zuiderveen Borgesius, 'Strengthening Legal Protection against Discrimination by Algorithms and Artificial Intelligence', *The International Journal of Human Rights* 24, no. 10 (25 November 2020): 1572–93, https://doi.org/10.1080/13642987.2020.1743976; Gerards and Zuiderveen Borgesius, 'Protected Grounds and the System of Non-Discrimination Law in the Context of Algorithmic Decision-Making and Artificial Intelligence', June 2022.

b) To assess and evaluate specific interpretations of equality and their response to the egalitarian risks algorithms present **(evaluation).**

c) To facilitate our ability to deepen our understanding of the egalitarian harms algorithms risk to impose, the conditions under which these risks can manifest, and the safeguards that are needed to address them **(navigation).**

Drawing inspiration from both European equality and non-discrimination law and political philosophical theories of justice and informed by the (practical) functioning of algorithmic decision-making systems and the particular challenges these systems bring along, I develop a multidimensional concept of equality. This model acts as the dissertation's analytical and evaluative point of departure and is used to a) identify the egalitarian harms algorithms risk to impose **(Part I)**, b) assess and evaluate the legal conceptualisation of equality **(Part II)** and c) locate and examine additional conceptualisations of equality to further deepen our understanding of the egalitarian harms algorithms can produce **(Part III)**. Following these three research exercises, I conclude the research project with a set of recommendations in order to strengthen the notion of equality as a regulatory instrument **(Part IV)**.

RESEARCH CONTRIBUTION - This dissertation offers a novel contribution to the on-going scholarly and policy debate concerning the injustice of algorithmic decision-making systems and does so through its primary focus on the notion of equality as a means for articulation and evaluation.

I assert that algorithmically guided decision-making systems, through their reliance on differentiation techniques, continuously expose decision-subjects to the risk of an egalitarian harm: they risk the incurrence of unjustified (in)equalities. I therefore argue that the (un)fairness and (in)justice of an algorithmically guided decision, and the (un)fairness and (in)justice of the differentiating acts, should be located, articulated and evaluated in light of the egalitarian harms these systems are likely to produce.[25] Egalitarian harms can take (at least) three (interrelated) shapes: they can be moral, socio-relational or distributional in nature.[26] These harms can be experienced by any person exposed to algorithmic decision-making processes, regardless of their current social position, and regardless of the decision-ground that was used. Whether and to what

[25] For the relationship that exists between the notions of fairness and justice, see also Chapter 1, Section 1.3: "Constructing Fairness."

[26] I took the notion "justice-relevant goods" from Christian Schemmel. Christian Schemmel, 'Distributive and Relational Equality':, *Politics, Philosophy & Economics*, 15 September 2011, 127, https://doi.org/10.1177/1470594X11416774.

extent these inequalities are recognised and addressed, will depend upon the specific interpretation given to equality by policy makers, regulators, or theorists.

By reconsidering the concept of equality in light of the novel challenges algorithmic decision-making systems have introduced, my goal is to strengthen regulatory initiatives that wish to take aim at these technological artefacts. Following an investigation into the egalitarian harms these systems produce, I investigate whether these risks can be addressed via the current legal approach towards equality. Finally, I deepen our understanding of these harms, and the conditions under which they can manifest, through specific interpretations of equality found in theories of justice. In doing so, I develop an understanding of equality that can guide the evaluation of algorithmically guided decision-making systems in a more holistic manner, from cradle to grave. In other words, the research project provides a methodology to operationalise equality for the evaluation and regulation of algorithmic decisions.

C. Research Methodology, Structure and Outline

In the following paragraphs, I further clarify my research approach and structure. I identify the main research objectives I set out to achieve within each of this dissertation's four core research Parts and their associated Chapters. In addition, I highlight the (normative) choices made regarding this research project's scope.

Part I - (In)equality in the Algorithmic Environment: Identifying The Risk of Egalitarian Harm.

In **Part I** of this dissertation, I contextualise the core research problem by offering the reader a clear understanding of the algorithmic environment and the key functions equality can hold therein. The research objectives for this foundational part are threefold.

THE ALGORITHMIC RESEARCH ENVIRONMENT - In a first step, I substantiate the claim that the introduction of algorithms for the automation of decision-making has brought along several challenges that justify this research project **(Chapter 1)**. Because the social value of equality is at the centre of my research, I demonstrate how algorithmically generated inequalities differ in terms of their form and scope, and as a consequence thereof, have the potential to actively restructure society. To that end, I identify and elaborate upon the unique properties data-driven analytics techniques exhibit by investigating their practical functioning, the purposes for which

they are used, and the actors who deploy them. The resulting picture of the algorithmic environment constitutes the research background onto which the notion of equality is applied and in which it is operationalised.

SCOPE - Within this dissertation, the ***notion algorithmically guided decision-making*** is used as a broad umbrella term that refers to any instance where human decision-makers rely upon the support or guidance of algorithmic techniques for taking specific decisions, including measures to govern, monitor, modify or otherwise regulate human actions and behaviour.[27]

Furthermore, I am concerned with the impact algorithmically guided decision-making processes have on people or ***decision-subjects***.[28] The notion of decision-subject is understood broadly. It not only concerns those persons or groups to whom a specific, material and individualized decision-making outcome is or will be applied. The notion also refers to any person or group that is liable to experience egalitarian harms that result from the choices and practices that inform and shape the development, design, functioning, and deployment of algorithmic structures and decisions, as well as the outcomes they produce.[29]

EQUALITY AS MULTIDIMENSIONAL CONCEPT - A society's commitment to equality represents its commitment to equalise the enjoyment of certain fundamental interests among its citizens. In a second step of this research project, my goal is to unearth the main (policy) ideals typically ascribed to equality as a social value underlying the structuring of society **(Chapter 2)**. Following an extensive literature study into various political philosophical theories of justice, the ethics of

---

[27] An overview of these techniques will be given in Chapter 1. For discussions on the notion "algorithmic regulation", see also: Mireille Hildebrandt, 'Algorithmic Regulation and the Rule of Law', *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences* 376, no. 2128 (13 September 2018): 20170355, https://doi.org/10.1098/rsta.2017.0355; Karen Yeung and Martin Lodge, eds., *Algorithmic Regulation* (Oxford, New York: Oxford University Press, 2019); Sylvie Delacroix, 'Beware of "Algorithmic Regulation"', SSRN Scholarly Paper (Rochester, NY: Social Science Research Network, 1 February 2019), https://doi.org/10.2139/ssrn.3327191.

[28] For instance, discovery techniques may also be used to analyse climate or astronomical data. Of course, these systems too can be beneficial or harmful to people depending on how the knowledge is used.

[29] This also includes the differentiation decision-making systems impose and the (in)equalities they generate. For instance, suppose a financial institution develops a credit application algorithm. This decision might exclude certain groups from being able to obtain a loan. Imagine this group is defined by people who are left-handed. This decision might not only affect left-handed decision-subjects who actually apply for the loan at that institution. If it would become known that a financial institution considers left-handed people less trustworthy with money, this information might also have a negative impact on the position of every left-handed person in society. How these harms can be generated, will be explained in the following Chapters.

discrimination, European equality and non-discrimination laws and legal theory, I present equality as a concept that can be specified alongside three (interrelated) dimensions or axes: a moral, socio-relational, and distributive dimension. Within this model equality is not merely seen as a procedural notion, but as a concept that reflects substantive policy objectives.[30] It provides a representative, albeit still abstract, vantage point to reflect upon, and shape, equality as a goal-oriented notion. Through their interpretation of these dimensions, a law, policy, or theory can give specific normative substance to the concept of equality. This specific interpretation or ***conceptualisation of equality*** determines what fundamental interests ought to be equalised and

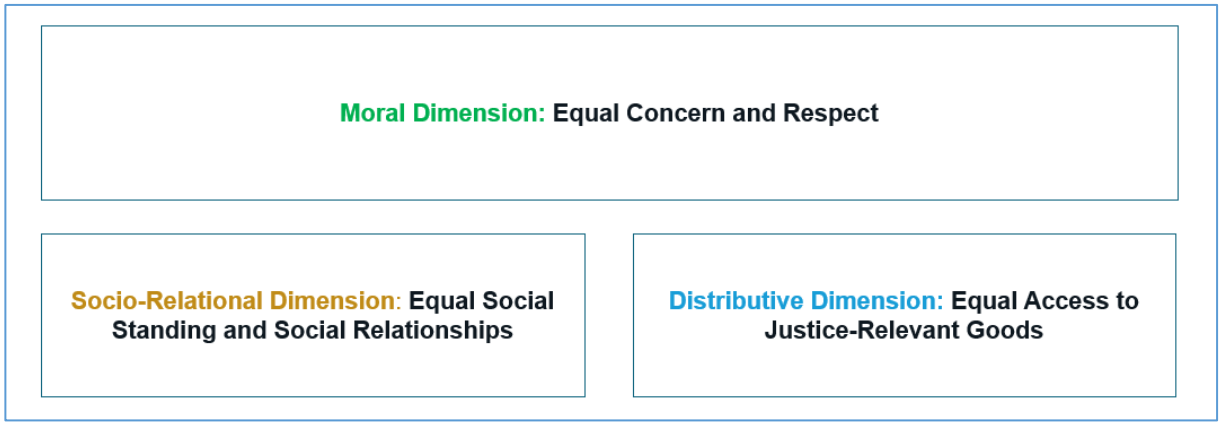

Figure 1. Equality as a multidimensional concept.

THE RISK OF ALGORITHMIC EGALITARIAN HARM – The third research objective is to specify the egalitarian harms algorithms can impose onto decision-subjects. In order to do so, I position the multidimensional model of equality against the algorithmic environment **(Chapter 2).** Following a dialogue between a functional understanding of algorithms on the one hand, and a theoretical, but goal-oriented understanding of equality on the other, I argue that algorithmic decision-making processes, through their functioning, have the potential to interfere with the distinct ideals ascribed to equality. These interferences represent ***a risk of egalitarian harm***. This risk arises from the differentiation algorithmic decisions impose and the inequalities they generate. The most specific ways in which algorithms impose this risk is by changing the social relationships people stand in (socio-relational inequality), and by hindering the desired allocation of certain justice-relevant goods (distributive equality).[31] Furthermore, via this identification function, I also demonstrate how these egalitarian risks are carried indiscriminately by each decision-subject regardless of their current social standing or make-up. **(Equality of Whom**).

[30] The equal or consistent application of a given procedure or rule.

[31] In trying to understand and define the core-research setting, and the associated injustices algorithms can impose, this research phase was also informed by real-life examples of discriminatory and inegalitarian algorithms challenged in news media, scholarship, or civil society.

THE FUNCTIONS OF THE MULTIDIMENSIONAL MODEL OF EQUALITY - Throughout this dissertation, the multidimensional model of equality is used as a practicable, analytical, and interpretative, i.e., normative lens, for the evaluation of algorithmically guided decision-making instances and structures, and the (regulatory) frameworks that govern their development and deployment. The model performs three specific purposes within this dissertation. As above-mentioned, I rely upon this understanding of equality to identify the egalitarian harms algorithms risk to impose **(Identification, Part I).** Once we know how algorithmic decisions risk to interfere with ideals of equality, the model can be used for evaluative and navigational purposes. If these egalitarian risks are considered to be critical, the specific conceptualisation of equality put forward by a given policy or law can be evaluated in light of their ability to identify and address said risks. I perform this exercise by uncovering and evaluating the legal conceptualisation of equality as defined by European equality and non-discrimination law **(Evaluation, Part II)**. Finally, I use the model as a navigational beacon to locate concrete conceptualisations of equality - which are chosen due to their correspondence with these dimensions - in order to deepen our understanding of algorithmic egalitarian harm, the conditions under which this harm can manifest, and the safeguards needed to avert this from happing **(Navigation, Part III).**

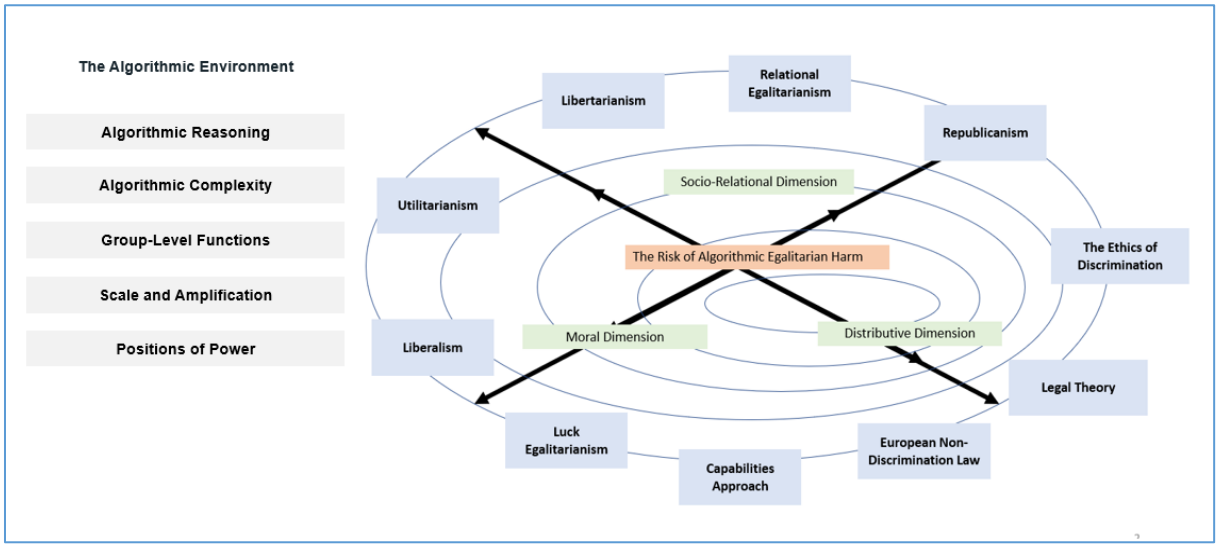

Figure 2. The Identification Function: The left-hand side represents the algorithmic background environment and its distinctive characteristics. The right-hand side depicts the research process, and sources that were consulted, underlying the formulation of equality as a multidimensional concept. In order to identify the egalitarian harm algorithms risk to impose, the multidimensional concept of equality is positioned against the algorithmic research environment.

Part II - Evaluating the Law: The Legal Conceptualisation of Equality and the Risk of Algorithmic Egalitarian Harm.

In Part II of this dissertation, I examine the legal conceptualisation of equality found within the Council of Europe (CoE) and European Union (EU) legal orders. The multidimensional model of equality plays a supportive role for the evaluation of the law. My main research objective in this Part is to identify the egalitarian objectives pursued by both systems, the egalitarian harms they guard citizens against, and the conditions under which those interests can be overridden. At the same time, I investigate the law's ability to address algorithmic injustice and demonstrate how the law's capacity to do so is informed (and constrained) by the (normative) choices made European lawmakers and the judiciary in their interpretation of equality.

First, I establish a methodology that allows me to uncover the legal conceptualisation of equality **(Chapter 3)**. Taking into account the risk of egalitarian harm identified within **Part I**, I argue that two axes of enquiry are of particular importance. In a first step, I investigate the law's (personal) range of protection. In a second step, I analyse to what extent acts of differentiation and inequality can be justified. The latter enquiry can be further specified. On the one hand, I explore the legal review procedures that are followed when the egalitarian interests of decision-subjects are weighed against the competing interests pursued by decision-makers. On the other hand, I assess to what extent the moral, socio-relational, or distributive egalitarian interests of decision-subjects are recognised as valid concerns when performing this review. I apply these axes of enquiry onto the CoE **(Chapter 4)** and EU's **(Chapter 5)** key equality provisions. In parallel, I assess the consequences of this approach for the regulation of algorithmically generated inequalities.

By concretising whom the law protects against which (algorithmic) inequalities in the pursuit of which ideals, and how, I evaluate the capabilities and limitations of the law in the face of algorithmic injustice. In subsequent order, should the legal framework be found to not adequately respond to the egalitarian challenges brought forth by algorithms, several options can be explored. We can formulate recommendations on how to amend the current legal approach. Or we can look for regulatory guidance elsewhere. For this, I turn to political philosophical theories of justice.

SCOPE – My analysis focuses on the conceptualisation of equality *in practice*, i.e., through (case-) law, rather than *in theory,* i.e. on the basis of academic accounts of what equality as a legal principle

should entail.[32] Hence, it excludes certain questions that are more practical in nature. For instance, on a procedural level, research has shown how the complexity and opacity of algorithmic decision-making systems has made it more difficult to prove the existence of discrimination. Similarly, legal scholars have argued that certain legal constructs, such as the conceptual distinction between direct and indirect discrimination, are difficult to maintain within the algorithmic environment.[33] Both examples highlight practical complexities in the application of equality and non-discrimination law: they tell us that the law might be unable to safeguard and enforce its substantive ambitions. In this research project, the main goal of the legal analysis is to uncover

[32] Certain scholars have proposed their own ideas as to what inequalities the law should address and on what grounds. These accounts can be used to highlight certain deficits of the law in practice, but they are not the main focus of the legal analysis. They have however co-informed the creation of the multidimensional model of equality. Tarunabh Khaitan, *A Theory of Discrimination Law*, *A Theory of Discrimination Law* (Oxford University Press), accessed 1 June 2020, http://www.oxfordscholarship.com/view/10.1093/acprof:oso/9780199656967.001.0001/acprof-9780199656967; Shreya Atrey, *Intersectional Discrimination* (Oxford: Oxford University Press, 2019), https://doi.org/10.1093/oso/9780198848950.001.0001; Iyiola Solanke, *Discrimination as Stigma: A Theory of Anti-Discrimination Law.* (Oxford, England ; Portland, Oregon: Hart Publishing, 2017). See also the many contributions in: Deborah Hellman and Sophia Moreau, eds., *Philosophical Foundations of Discrimination Law*, Philosophical Foundations of Law (Oxford: Oxford University Press, 2013), https://doi.org/10.1093/acprof:oso/9780199664313.001.0001; Kasper Lippert-Rasmussen, ed., *The Routledge Handbook of the Ethics of Discrimination*, 1st ed., Routledge Handbooks (New York: Routledge, 2018), https:///bookshop/product/The-Routledge-Handbook-of-the-Ethics-of-Discrimination-by-Kasper-Lippert-Rasmussen-editor/9781317400769.

[33] An analysis of the legal tools offered by non-discrimination laws for dealing with algorithmic discrimination, and an investigation as to whether the concepts utilised by the law are adapted to the reality of the algorithmic environment have already been conducted by others, including myself. See for example: Janneke Gerards and Frederik Zuiderveen Borgesius, 'Protected Grounds and the System of Non-Discrimination Law in the Context of Algorithmic Decision-Making and Artificial Intelligence', SSRN Scholarly Paper (Rochester, NY: Social Science Research Network, 2 November 2020), https://papers.ssrn.com/abstract=3723873; Janneke Gerards et al., *Algorithmic Discrimination in Europe: Challenges and Opportunities for Gender Equality and Non Discrimination Law* (LU: Publications Office of the European Union, 2021), https://data.europa.eu/doi/10.2838/544956; Gerards et al.; Raphaële Xenidis and Linda Senden, 'EU Non-Discrimination Law in the Era of Artificial Intelligence : Mapping the Challenges of Algorithmic Discrimination', in *General Principles of EU Law and the EU Digital Order*, ed. Ulf Bernitz et al. (Kluwer Law International, 2020), 151–82, https://cadmus.eui.eu//handle/1814/65845; Borgesius, 'Strengthening Legal Protection against Discrimination by Algorithms and Artificial Intelligence'; Leese, 'The New Profiling: Algorithms, Black Boxes, and the Failure of Anti-Discriminatory Safeguards in the European Union'; Anna Lauren Hoffmann, 'Where Fairness Fails: Data, Algorithms, and the Limits of Antidiscrimination Discourse', *Information, Communication & Society* 22, no. 7 (7 June 2019): 900–915, https://doi.org/10.1080/1369118X.2019.1573912; Naudts, 'How Machine Learning Generates Unfair Inequalities and How Data Protection Instruments May Help in Mitigating Them'; Sandra Wachter, 'Affinity Profiling and Discrimination By Association in Online Behavioral Advertising', 2021, https://doi.org/10.15779/Z38JS9H82M; Wachter, Mittelstadt, and Russell, 'Why Fairness Cannot Be Automated'.

what those ambitions are in the first place and how they relate to the injustices algorithms risk to produce.

Regarding the difference between the notions discrimination and differential treatment, the following distinction will be maintained. I refer to "discrimination" as forms of differential treatment that are considered problematic from a legal perspective, i.e., those forms of differentiation that are considered unjustified under equality and non-discrimination laws. In case of uncertainty regarding whether or not a given treatment constitutes discrimination in the legal sense, the neutral notion differential treatment will be used. In general, however, the use of more neutral terminology will be preferred in my discussion of the research problem. Hence, I will try to maintain the distinction between justified and unjustified differentiation and differential treatment, justified and unjustified (in)equality.[34]

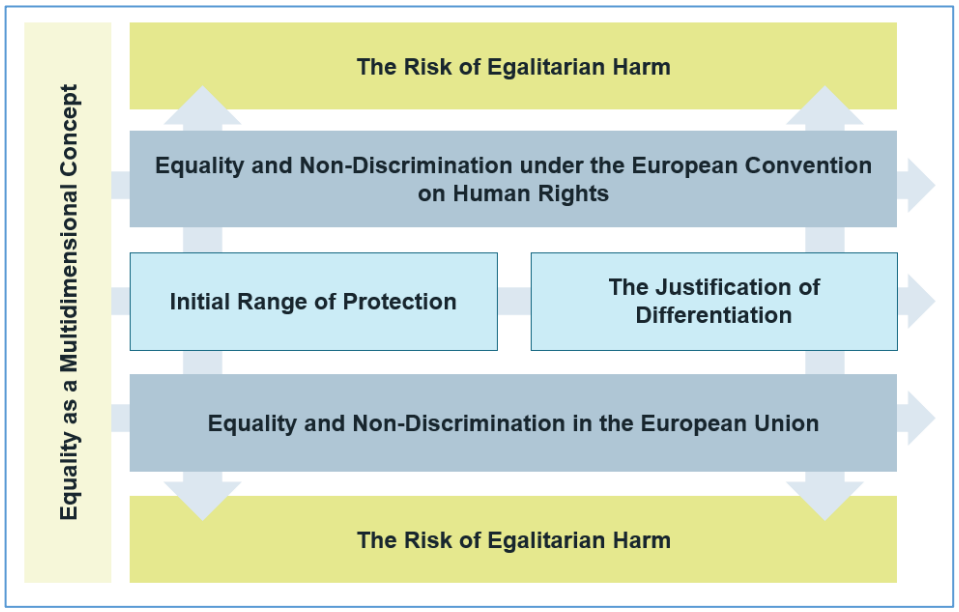

Figure 3 The Evaluative Function: With the support of the multidimensional concept of equality, I uncover and evaluate legal conceptualisation of equality in order to assess the law's to ability to address the egalitarian harms algorithm risk to impose.

[34] For a more elaborate explanation on the relationship between equality and non-discrimination, see also Chapter 2, Section 2.3 "Algorithmic Inequality: Related Notions."

Part III - Navigating Political Philosophy: Socio-Relational and Distributive Conceptualisations of Equality and the Articulation of Algorithmic Egalitarian Harm.

Theories of justice map the interplay and alignment between the legal, economic, and civil aspects of social life. They propose principles regarding the desired structure of society so it best accommodates the interests of citizens?[35] Within these theories, equality is often given great emphasis as a structuring principle.

I**n Part III** of this dissertation, I make use of the model's navigational function. Following my literature study, I have located specific socio-relational and distributive interpretations of equality found in theories of justice. These notions are transposed onto algorithmic environment. They have been chosen due to their correspondence with the multidimensional model of equality. In an effort to better articulate the socio-relational harms algorithms risk to impose, the notions of domination and oppression are examined **(Chapter 6).** Finally, I conclude this research analysis, I look into two notions that are (primarily) distributive in nature, namely equality of primary goods and equality of capability **(Chapter 7).**

Within the overall research framework, these interpretations of equality play a twofold role. Each conceptualisation can be seen as an alternative to the law. Like the normative rationales that ground the legal conceptualisation of equality, these political philosophical notions too can be used deepen our understanding of the egalitarian harms people risk incurring, the conditions under which such harms can manifest, and the institutional safeguards and resources decision-subjects need to strengthen their position vis-à-vis decision-makers. At the same time, we can investigate to what extent the law can be complemented by this deeper understanding in order to establish a more holistic and robust response towards the regulation of algorithmic decision-making. This seems particularly useful if we are to discover in Part II that the law's approach towards algorithmic injustice is insufficient.

I made the choice to focus on specific and concrete notions of equality, rather than specific philosophical traditions. The views I apply onto the algorithmic environment can moreover be

[35] Philip Pettit, *Judging Justice. An Introduction to Contemporary Political Philosophy* (London, Boston and Henley: Routlegde & Kegan Paul, 1980), 23–26. Pettit notes how theories of justice form the main subject of investigation for the political philosopher. The political philosopher will evaluate and compare these theorised versions of society in an effort to seek an optimum, i.e., to seek justice.

tied to specific authors: Philip Pettit and Iris Marion Young (domination and oppression), John Rawls (equality of primary goods) and Martha Nussbaum (equality of capability). If one believes that the risks identified in Part I should be on a regulators' radar, these conceptions of equality may function as an appropriate starting point to reflect upon a set of (alternative) conditions regarding the regulation of algorithmically guided decision-making systems. A point of caution is nonetheless in order: the chosen socio-relational and distributive conceptions of equality are not necessarily mutually compatible. Though I highlight synergies between them, more research is necessary to verify whether a consistent and coherent can be formulated on the basis of the aforementioned notions. Moreover, alternative conceptualisations of equality exist. Finally, certain theories of justice do not take equality as their vantage point. For example, some focus on utility or preferences as a metric of justice (utilitarianism), whereas others emphasise the importance of individual self-ownership (libertarianism or neoliberalism). These too could be worth exploring in future research projects.

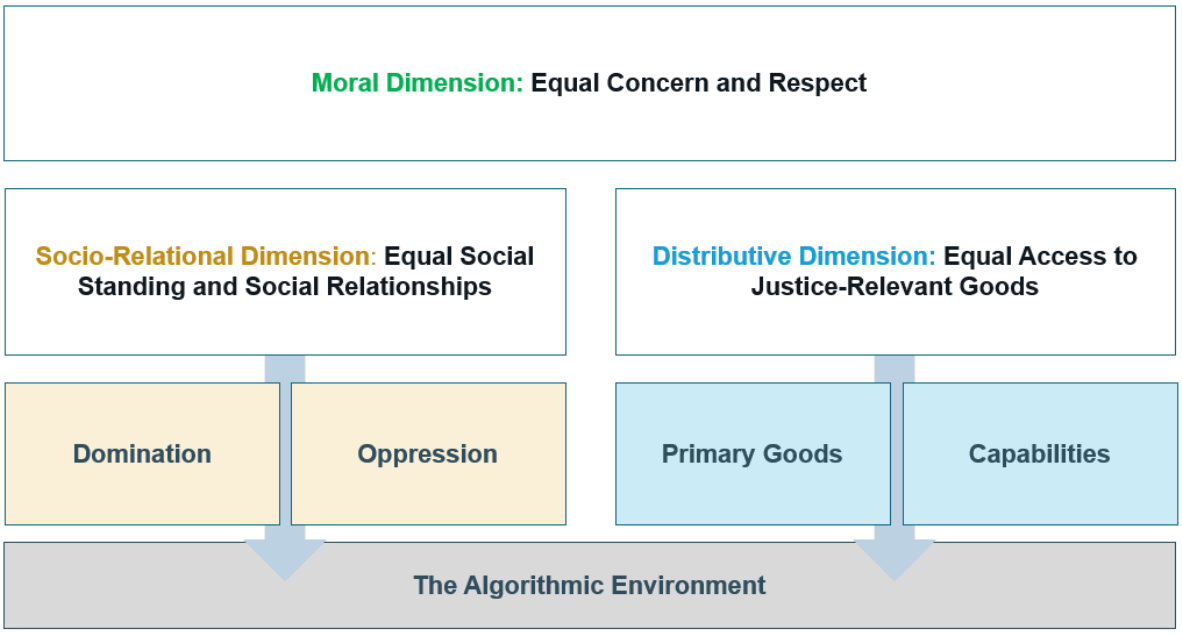

Figure 4. The Navigational Function: Using the multidimensional concept as a navigational beacon, various socio-relational and distributions notions of equality are identified and transposed onto the algorithmic environment to concretise the egalitarian harms algorithms risk to produce, the conditions under which these harms may manifest in practice, and the institutional safeguards that can be provided to decision-subjects and society at large to strengthen their position in an algorithmically mediated environment.

Part IV – Reconsidering the Concept of Equality for the Evaluation and Regulation of Algorithmically Guided Decision-Making: Concluding Remarks and Recommendations.

This dissertation's concluding sections summarise the main research findings **(Chapter 8).** Following the identification, articulation, evaluation, and navigation exercises performed within this research project, I highlight synergies identified between justice and the law, and concretise how the latter (or components thereof) can be informed by the former (and vice versa). I formulate a set of normative recommendations that can strengthen the future regulation and governance of algorithmically guided decision-making. By identifying gaps within the current regulatory approach, I formulate suggestions on how the position of equality as a social value can be strengthened when it is placed against competing interests. In addition, I argue that egalitarian interests of decision-subjects should find support through a robust set of institutional safeguards. In this regard, the research findings can help substantiate the importance of current safeguards instituted by the law, while at the same time highlight areas for improvement.

METHODOLOGICAL LIMITATIONS - In recent years, the European Union has launched several initiatives, both extra-legal, such as the HLEG on AI's Guidelines on Trustworthy AI,[36] and legal, such as the proposed AI Regulation.[37] Similarly, the Council of Europe's Ad Hoc Committee on AI (CAHAI) explored the need and potential for a legal framework design, development and application of AI.[38] These regulatory initiatives highlight equality and non-discrimination as

[36] The High-Level Expert Group on AI's guidelines towards trustworthy AI note how AI, as an active driver towards change, could "foster equality in the distribution of economic, social and political opportunity." At the same time, the guidelines' underlying human-rights based approach cautions against the use of AI in ways that interfere with the fundamental right to equality and non-discrimination. Similarly, the principle of fairness they advance should ensure groups and individuals are free from unfair bias, discrimination and stigmatisation. High-Level Expert Group on Artificial Intelligence, 'Ethics Guidelines for Trustworthy AI', 9, 10 and 12.

[37] Proposal for a Regulation of the European Parliament and of the Council laying down harmonised rules on Artificial Intelligence (Artificial Intelligence Act) and amending certain Union Legislative Acts. COM(2021) 206 Final, (Proposed AI Act). The explanatory memorandum notes (Section 3.5). In the proposal itself, recital 15 mentions how certain AI technologies can cause serious harm and should be prohibited because they contradict Union values, including equality, and Union fundamental rights, including the right to non-discrimination. Other references to equality or non-discrimination can be found in recitals 17, 28, 33, 35, 36, 37, 38, 39, 44, 45 and 47.

[38] See for instance: Ad Hoc Committee on Artificial Intelligence (CAHAI). 'Feasibility Study'. Council of Europe, 17 December 2020, available via: https://rm.coe.int/cahai-2020-23-final-eng-feasibility-study-/1680a0c6da.; Ad Hoc Committee on Artificial Intelligence (CAHAI). 'Possible elements of a legal framework on artificial intelligence, based on the Council of Europe's standards on human rights, democracy and the rule of law', 3 December 2021; available via:

important guiding principles in the development and deployment of algorithmically guided decision-making processes. Although these frameworks will not be the subject of my research, they are testimony to the importance of this research project: despite its prominent position in emerging regulatory initiatives, the notion of equality often remains undefined. This lack of precision could limit their ability to fully capture algorithmic egalitarian harms.

I argue that an examination into the function equality holds in both legal traditions goes a long way in clarifying the evaluative and guiding role this value can have for the future governance of algorithmic decision-making practices. It seems only reasonable that for the interpretation of the notions of equality and non-discrimination within these regulatory efforts, and in order to ensure legal consistency, reference can be made to their interpretation within the legal frameworks I do discuss. At the same time, regulators and algorithmic decision-makers can further deepen and strengthen their understanding of equality by turning to political philosophy. Political philosophical notions of equality can either be an alternative or complement to the law. Through an analysis of both the law and political philosophy, this research project therefore makes an active contribution to the future governance of algorithmically guided decision-making processes.

D. A Word on the Notion of Fairness

The European High Level Expert Group on AI's guidelines on trustworthy AI solidified fairness and justice as leading principles in the algorithmic governance debate.[39] In their view, fairness, justice, and equality share a close relationship: *"The development, deployment and use of AI systems must be fair. While we acknowledge that there are many different interpretations of fairness, we believe that fairness has both a substantive and a procedural dimension. The substantive dimension implies a commitment to: ensuring equal and just distribution of both benefits and costs, and ensuring that individuals and groups are free from unfair bias, discrimination and stigmatisation […] The procedural dimension of fairness entails the ability to contest and seek effective redress*

https://rm.coe.int/cahai-2021-09rev-elements/1680a6d90d.; Isaac Ben-Israel et al., 'Towards Regulation of AI Systems: Global Perspectives on the Development of a Legal Framework on Artificial Intelligence Systems Based on the Council of Europe's Standards on Human Rights, Democracy and the Rule of Law', Council of Europe Study (Council of Europe, December 2020), https://rm.coe.int/prems-107320-gbr-2018-compli-cahai-couv-texte-a4-bat-web/1680a0c17a.

[39] High Level Expert Group on AI, 'Draft Ethics Guidelines for Trustworthy AI', Text, Shaping Europe's digital future - European Commission, 18 December 2018, 10, https://ec.europa.eu/digital-single-market/en/news/draft-ethics-guidelines-trustworthy-ai.

against decisions made by AI systems and by the humans operating them."[40] In the group's self-assessment list towards trustworthy AI, fairness has been defined as a notion that refers to "a variety of ideas known as equity, impartiality, egalitarianism, non-discrimination and justice."[41]

In this dissertation, I shall take a similar view towards the connection between equality, justice, and fairness. More specifically, equality is seen as a value that contributes to the fairness of a given decision. Equality alone does not suffice for a decision to be considered fair though. Indeed, following the abovementioned definition, whether a decision is fair, will also depend upon other factors. For instance, whether the decision is accompanied by mechanisms of accountability and transparency, which can be seen as a precondition to ensure effective redress.

The question whether a given act of differentiation was fair can therefore not be explained through equality alone. Because this dissertation focuses on the concept of equality however, I will often make use of the notions "justified or unjustified", rather than "fair or unfair" to categorise a given act of differentiation and the inequalities that follow thereout.

Where an act of differentiation cannot be justified due to the inequalities it produces, it will work against the overall fairness of a given decision-making procedure. In this sense, unjustified differentiation is quasi-synonymous with unfair differentiation (and vice versa). At the same time, fairness acts as the broader principle. As a consequence, the absence of appropriate fairness standards can have an impact on the evaluation of the inequalities a given decision produces. For instance, where decision-making procedures are entirely opaque and closed off from a (democratic) review, it is difficult to see how the inequalities they impose onto people can be justified.

---

[40] Alongside respect for human autonomy, prevention of harm and explicability. The initial draft carried its inspiration more clearly on its sleeve. In that draft, the principles were: the principle of beneficence ("Do Good"); the principle of non-maleficence ("Do no Harm"); the principle of autonomy ("Preserve Human Agency"); the principle of justice ("Be Fair') and the principle of explicability. In the HLEG AI's initial draft, justice was recognised as the requirement to be fair. The guidelines were greatly influenced by Beauchamp and Childress' influential account on biomedical ethics, where justice, alongside respect for autonomy, beneficence and non-maleficence, is posited as one of the four principles of biomedical ethics. Under their model, justice requires *amongst others* the *fair* distribution of risks, burdens and costs. 'Principles of Biomedical Ethics - Paperback - Tom L. Beauchamp; James F. Childress - Oxford University Press', accessed 13 July 2020, https://global.oup.com/ushe/product/principles-of-biomedical-ethics-9780190640873?cc=be&lang=en&..: High-Level Expert Group on Artificial Intelligence, 'Ethics Guidelines for Trustworthy AI', 12. High Level Expert Group on AI, 'Draft Ethics Guidelines for Trustworthy AI', 10.

[41] High Level Expert Group on AI, 'The Assessment List for Trustworthy Artificial Intelligence (ALTAI) for Self-Assessment', 17 July 2020, 27.

PART I – (In)equality in the Algorithmic Environment: Identifying The Risk of Egalitarian Harm.

1 Algorithmically Guided Decision-Making: Differentiation and Inequality

The main research objective of this dissertation is to operationalise the concept of equality as a guiding mechanism for the identification, articulation and evaluation of unjustified egalitarian harm imposed by these decision-making systems, and the regulatory response given thereto. If we want to reposition the notion of equality with this purpose in mind, we must first understand the research environment in which it will be applied. In this Chapter, I therefore want to provide the reader with a better understanding of algorithmic decision-making systems and the social challenges they introduce.

To that end, I first describe the technical functions algorithms perform and the purposes for which they are used. I explain how these systems rely upon differentiation, and in doing so, produce (in)equality **(Section 1.1)**. Through an exploration of their technical functioning, and the particular advantages these technologies offer decision-makers, I then demonstrate how the introduction of algorithms within decision-making processes initiates social change as a result of the inequalities algorithms impose and produce. I argue however that algorithmic systems possess a series of properties that render the inequalities they generate distinctive in terms of their form and scope. Due to these unique properties; algorithmic inequalities have the potential to restructure the fabric of society alongside new and existing dimensions: they may not only reinforce existing social injustice, but they may also introduce new forms of (non-representational) injustice. As a consequence, the process of social change they initiate, may challenge the way we approach equality as a value underlying the structuring of society **(Section 1.2)**. In the final Section, I contend that this restructuring capacity should be examined in relation to the positions of power that are held by those actors who control the means of (algorithmic) knowledge discovery, production, interpretation, and application **(Section 1.3)**.

1.1 The Functions of Algorithms

Though extensive discussions can be held regarding the actual, correct, and accurate definition of algorithms, I will make use of the notion "algorithmically guided decision-making process" as it has come to be known in contemporary media, legal scholarship and the everyday. That is, as a broad, catch-all term for decision-making procedures aided through a variety of technical and

often knowledge or data-driven techniques.[42] In order to better understand the subject of my analysis however, a brief explanation regarding the nature and functioning of algorithms and algorithmically guided decision-making systems seems appropriate.[43]

1.1.1 Algorithms

ALGORITHMS - Algorithms can be defined as an "unambiguous procedure to solve a problem or class of problems, […] typically composed of a set of instructions or rules that take some input data and return outputs."[44] The notion "algorithm" refers to a set of techniques that can be used for a variety of purposes: to discover hidden patterns and correlations within data sets; to categorise, classify, prioritise or filter data or information; to look for group-level relationships between (groups of) individuals which can be applied as a knowledge basis to differentiate among them; to automate and streamline a given procedure or decision-making process that was previously analogue; etc.[45] Algorithms can be entirely hand-coded by a designer, or they can be generated from data.[46] A typical distinction is often made between rule-based and learning algorithms.[47]

RULE-BASED ALGORITHMS – Rule-based algorithms consist of "a set of assertions and a set of rules" which define how a system should act in a particular situation. They can typically be expressed as

---

[42] Rob Kitchin, 'Thinking Critically about and Researching Algorithms', *Information, Communication & Society* 20, no. 1 (2017): 14–29, https://doi.org/10.1080/1369118X.2016.1154087.

[43] As explained in the introductory chapter, the notion of algorithmically guided decision-making will refer to any decision-making instance where a human decision-maker relies upon the support or guidance of the below-mentioned algorithmic techniques for taking a specific decision. The notion decision also includes measures that are used to govern, monitor, modify or otherwise regulate human actions and behaviour. I am concerned with the impact these decisions have on people or decision-subjects. This notion will refer to any person or group that is liable to experience an impact, and *in casu* an egalitarian harm, that stems from the choices and practices that have informed the development, design, functioning, and deployment of algorithmic structures and decisions, including the outcomes they produce. In other words, the notion not only captures people that are the direct target of an algorithmic decision-making practice, but also those persons that can experience an indirect impact due to their development and deployment.

[44] Castelluccia and Le Metayer, 'Understanding Algorithmic Decision-Making: Opportunities and Challenges', 3.

[45] Diakopoulos, 'Accountability in Algorithmic Decision Making'.

[46] Castelluccia and Le Metayer, 'Understanding Algorithmic Decision-Making: Opportunities and Challenges', 3.

[47] For more information on these distinctions, see for example: Hildebrandt, 'Algorithmic Regulation and the Rule of Law'; Harry Surden, 'Artificial Intelligence and Law: An Overview', *Georgia State University Law Review* 35, no. 4 (1 June 2019), https://readingroom.law.gsu.edu/gsulr/vol35/iss4/8. See also: Kulk and van Deursen's research report on algorithmically guided decision-making which helped structure this section: Stefan Kulk and Stijn van Deursen, *Juridische Aspecten van Algoritmen Die Besluiten Nemen: Een Verkennend Onderzoek.* (Boom Juridisch, 2020).

a set of if/then statements.[48] Due to their underlying formal logic, rule-based systems appear highly suitable for the automation of decision-trees or other decisions where knowledge can be easily translated into rules.[49] For example, they have been used to streamline and standardise certain administrative procedures.[50] Because they follow a fixed and streamlined structure, rule-based systems are considered efficient, consistent, deterministic, predictable, explainable and retraceable.[51] However, their rigidity comes at a cost as not every problem can be moulded into an if/then structure. Hence, these systems are less appropriate in their support of decisions that require discretion, subjective interpretation, or contextual considerations.[52]

MACHINE LEARNING ALGORITHMS – Making use of enormous amounts of data, machine learning algorithms have the capacity to automatically learn and in so doing, improve and adapt their functioning.[53] The more the machine can learn, the better and more efficient it should become in the performance of its tasks. The learning process involves the training, testing, and validation of the algorithm through data. This learning phase can be supervised, unsupervised or reinforced.[54] In practice, a decision-making process may be guided by a variety of learning techniques, whereby each technique performs a particular function towards the achievement of the desired end-result. Developments in machine learning and the increased availability of data are often seen as the catalysts of the current artificial intelligence revolution.[55]

---

[48] Crina Grosan and Ajith Abraham, *Intelligent Systems*, vol. 17, Intelligent Systems Reference Library (Berlin, Heidelberg: Springer Berlin Heidelberg, 2011), 149, https://doi.org/10.1007/978-3-642-21004-4. See also: Han Liu, Alexander Gegov, and Mihaela Cocea, *Rule Based Systems for Big Data*, vol. 13, Studies in Big Data (Cham: Springer International Publishing, 2016), 2, https://doi.org/10.1007/978-3-319-23696-4.

[49] Kulk and van Deursen, *Juridische Aspecten van Algoritmen Die Besluiten Nemen: Een Verkennend Onderzoek.*, 36–37.

[50] Hildebrandt, 'Algorithmic Regulation and the Rule of Law', 3.

[51] Kulk and van Deursen, *Juridische Aspecten van Algoritmen Die Besluiten Nemen: Een Verkennend Onderzoek.*, 37. Hildebrandt, 'Algorithmic Regulation and the Rule of Law', 2; Grosan and Abraham, *Intelligent Systems*, 17:175.

[52] See for a full overview of drawbacks also: Kulk and van Deursen, *Juridische Aspecten van Algoritmen Die Besluiten Nemen: Een Verkennend Onderzoek.*, 37. Grosan and Abraham, *Intelligent Systems*, 17:176.

[53] Whereas learning algorithms are typically referred to as *data-driven*, rule-based algorithms can nonetheless depend upon input data too. See also: Grosan and Abraham, *Intelligent Systems*, vol. 17, chaps 10, Machine Learning. Kulk and van Deursen, *Juridische Aspecten van Algoritmen Die Besluiten Nemen: Een Verkennend Onderzoek.*; Gerards et al., *Algorithmic Discrimination in Europe*.

[54] Solon Barocas, Moritz Hardt, and Arvind Narayanan, 'Fairness in Machine Learning', n.d., 127; Liu, Gegov, and Cocea, *Rule Based Systems for Big Data*; Grosan and Abraham, *Intelligent Systems*.

[55] For a more nuanced view, see also: Michael I. Jordan, 'Artificial Intelligence—The Revolution Hasn't Happened Yet', *Harvard Data Science Review* 1, no. 1 (1 July 2019), https://doi.org/10.1162/99608f92.f06c6e61.

Supervised learning involves training data that is *labelled*. This means that the algorithm has access to examples of which the desired category, class or outcome is already known. It must then learn from these known examples to reproduce and apply the correct outcome for instances that are currently unknown or unlabelled. A typical supervised learning technique is classification, whereby an algorithm is trained to detect to which class or category given piece of data belongs.[56] For example, using the payment history of previous clients, a financial institution may train an algorithm to predict the creditworthiness of new clients.

Unsupervised learning techniques do not depend upon labelled data. Instead, unsupervised algorithms are given a series of instructions to explore data and are tasked to discover hitherto unknown patterns and correlations. Examples of unsupervised learning techniques are association (i.e., finding correlations between data) and clustering (i.e., grouping data into homogenous segments).[57] Such techniques may be used to learn more about people's behaviours and preferences. For instance, a music platform may deploy an association technique to find an association or connection between people's listening behaviour and taste in music by identifying those songs that were typically listened to during user sessions. Similarly, a human resource agency may look to cluster together client firms based on the hiring criteria the latter use so they can more efficiently allocate job applicants to these clients.[58] Even though unsupervised algorithms are typically trusted to perform their tasks more autonomously, the system's output can be made subject to a periodic and systemic review, for instance to verify whether the desired results have been obtained.[59] Algorithms learn through reinforcement when they are given positive or negative feedback depending on whether they have achieved a desired aim or not. Reinforcement learning has been used for real-time bidding and online advertising purposes, where the algorithm is instructed to find the ad with the highest click-through rate.[60]

---

[56] Gerards et al., *Algorithmic Discrimination in Europe*, 35.

[57] Liu, Gegov, and Cocea, *Rule Based Systems for Big Data*, 13:2; Grosan and Abraham, *Intelligent Systems*, 17:265.

[58] Vasilis Verroios et al., 'Client Clustering for Hiring Modeling in Work Marketplaces', in *Proceedings of the 21th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, KDD '15 (New York, NY, USA: Association for Computing Machinery, 2015), 2187–96, https://doi.org/10.1145/2783258.2788589.

[59] Kulk and van Deursen, *Juridische Aspecten van Algoritmen Die Besluiten Nemen: Een Verkennend Onderzoek.*, 41.

[60] Han Cai et al., 'Real-Time Bidding by Reinforcement Learning in Display Advertising', in *Proceedings of the Tenth ACM International Conference on Web Search and Data Mining*, 2017, 661–70, https://doi.org/10.1145/3018661.3018702; 'How Reinforcement Learning Chooses the Ads You See', *VentureBeat* (blog), 23 February 2021, https://venturebeat.com/2021/02/23/how-reinforcement-learning-chooses-the-ads-you-see/. There are also more complex learning techniques, inspired by nature, such as (deep) neural networks and evolutionary learning. These techniques are mainly associated with more intuitive or sensory tasks that are not easily translated into formal, rule-based structures, such as image and sound recognition or natural language processing. The main idea that drives nature-

Finally, the notion of data mining has been used for the deployment of data-driven techniques for the discovery and extraction of knowledge and patterns from data that were previously unknown.[61] It should be observed that learning algorithms discover correlations, rather than causal links between data points.[62] Even though a causal link may not be present, and given sufficient data, correlations are often seen as a reliable enough to function as a basis of decision-making.[63]

THE IMPORTANCE OF DATA - Most algorithmic processes derive their use and purpose through data. As they require training and are moreover expected to improve their functioning over time, machine learning algorithms are highly dependent upon the quality of the available data.[64] Supervised learning techniques require data that is labelled. Consequently, they may be less flexible and adaptive within highly dynamic environments where contextual changes need to be taken into account quickly.[65] Though unsupervised learning and deep-learning techniques are generally considered more adaptive, their adaptability comes at the cost of explicability. It can be difficult to explain how a system works and adapts, and once deployed, end-users may have little control over how the system's functioning will evolve.[66]

KNOWLEDGE CREATION AND APPLICATION - Algorithmic techniques are used for either one of two broad purposes, or a combination thereof: knowledge creation and knowledge application.[67] On

---

inspired systems is the limited capacity of formal structures for certain types of problem solving. For instance, mathematical or logical functions are not relied upon in determining the face of a friend. Hence, when we let computers perform the task of facial recognition, the application of a formal reasoning process to do so seems less appropriate. Instead, the computer draws inspiration from the functioning of the brain in order to replicate the intuitive process of facial recognition. They are the epitome of black-box algorithms. John Fulcher, 'Computational Intelligence: An Introduction', *Computational Intelligence*, n.d., 18. See also: Kulk and van Deursen, *Juridische Aspecten van Algoritmen Die Besluiten Nemen: Een Verkennend Onderzoek.*, 41; Gerards et al., *Algorithmic Discrimination in Europe*, 36.

[61] Liu, Gegov, and Cocea, *Rule Based Systems for Big Data*, 13:vii. Liu, Gegov, and Cocea, 13:2.

[62] Kulk and van Deursen, *Juridische Aspecten van Algoritmen Die Besluiten Nemen: Een Verkennend Onderzoek.*, 39.

[63] Brent Daniel Mittelstadt et al., 'The Ethics of Algorithms: Mapping the Debate', *Big Data & Society* 3, no. 2 (December 2016): 5, https://doi.org/10.1177/2053951716679679.

[64] In the case of supervised learning techniques, the labels too must be reliable and qualitative. Kulk and van Deursen, *Juridische Aspecten van Algoritmen Die Besluiten Nemen: Een Verkennend Onderzoek.*, 40. Surden, 'Artificial Intelligence and Law', 1315–16.

[65] Gerards et al., *Algorithmic Discrimination in Europe*, 38.

[66] Gerards et al., *Algorithmic Discrimination in Europe*. In reference to : Max Vetzo, J. H. Gerards, and Remco Nehmelman, *Algoritmes en grondrechten*, Montaigne Centrum 10 (Den Haag: Boom juridisch, 2018), 49.

[67] Castelluccia and Le Metayer define three specific classes of algorithmically guided decision-making systems: a) systems aimed at improving general knowledge or technology, b) systems aimed at improving or developing new digital

the one hand, a variety of these techniques have as their core function the expansion of human knowledge. They look to discover hitherto unknown, yet relevant information, relations, patterns, correlations that up until now went undiscovered. This also includes undiscovered knowledge about people, including group-level relationships.[68] On the other hand, algorithmic techniques can be used to facilitate or streamline the application of knowledge. In the latter case, algorithms may give human decision-makers greater flexibility and range in the development and application of traditional, yet complex decision-support instruments, such as surveys and questionnaires, which may have been developed through non-algorithmically guided psychological and sociological research projects. Because the analysis and application of such (analogue) instruments can be facilitated through automation, the way they are developed may face less limits concerning their scope and scale.[69] Within a given context, algorithms can be used to perform both functions. For instance, recommender systems uncover the interests of end-users, and then use this new-found knowledge to decide upon what content would best suit their preferences.

Algorithmically guided decision-making systems can vary in their degree of complexity, automation and human oversight and presence. Certain algorithmically guided decision-making processes will follow a fixed and deterministic ruleset, others will adapt and learn over time. Some will operate in ways that are fully automated and autonomous, while others will require human supervision or assistance.[70] The decision-maker may rely upon one technique or a combination thereof. The tasks of the algorithm can be descriptive (analyse what will happen), diagnostic (clarify what has happened), predictive (predict what will happen) or prescriptive (give advice on what should happen).[71] Algorithms can be used throughout the entire decision-chain, or only to support a specific component thereof.

Regardless of their degree of automation and complexity, these systems always come about through the deliberate choice and actions of the people who build or deploy them, or find their

---

services, and c) systems that are to be integrated int cyber physical systems, like self-driving cars or care robots. Castelluccia and Le Metayer, 'Understanding Algorithmic Decision-Making: Opportunities and Challenges', 5.

[68] Salome Viljoen, 'Democratic Data: A Relational Theory For Data Governance', SSRN Scholarly Paper (Rochester, NY: Social Science Research Network, 11 November 2020), https://doi.org/10.2139/ssrn.3727562.

[69] There would be less worry for these questionnaires to impose an administrative burden on those people who typically had to read and analyse them.

[70] John Danaher, 'Freedom in an Age of Algocracy', The Oxford Handbook of Philosophy of Technology, 10 May 2022, https://doi.org/10.1093/oxfordhb/9780190851187.013.16. See also: Reuben Binns and Michael Veale, 'Is That Your Final Decision? Multi-Stage Profiling, Selective Effects, and Article 22 of the GDPR', *International Data Privacy Law* 11, no. 4 (1 November 2021): 319–32, https://doi.org/10.1093/idpl/ipab020.

[71] Kulk and van Deursen, *Juridische Aspecten van Algoritmen Die Besluiten Nemen: Een Verkennend Onderzoek.*, 35–36.

use in solving a specific problem appropriate.[72] The decisions that enter into their design and operation are value-laden, and the values in question have their roots either in a specific person, such as the designer or programmer, society at large (reflected in the data), or both.[73] Even though algorithms cannot be decoupled from the people that have built them or will interact with them upon deployment, it would be wrong to refer to algorithms as being passive artefacts. Like most technologies, algorithms have an active or transformative influence. Algorithmic technologies have become an integral component of our wider socio-technical living environment, and as will be demonstrated in Section 1.2, through their operations, they can both take and shape meaning, not only in expected, but also unexpected and unintended ways.[74]

1.1.2 Algorithmic Differentiation and (In)equality

This dissertation's title raises the question whether and when differentiation is unfair? As part of most decision-making procedures, and at various moments therein, people are made subject to instances of differentiation or differential treatment; instances where, throughout the various stages of decision-making - from data collection to deployment - distinctions have been imposed between or onto people with a concrete impact. This differentiating nature of algorithmically guided decision-making processes has raised several concerns regarding the justice and fairness of these systems, and more specifically how societal values, like equality can be guaranteed.[75] At

---

[72] Annette Zimmermann, 'The Power of Choosing Not to Build: Justice, Non-Deployment, and the Purpose of AI Optimization', 2021.

[73] Mittelstadt et al., 'The Ethics of Algorithms: Mapping the Debate'; Peter-Paul Verbeek, 'Materializing Morality: Design Ethics and Technological Mediation', *Science, Technology, & Human Values* 31, no. 3 (May 2006): 361–80, https://doi.org/10.1177/0162243905285847; Mike Ananny, 'Toward an Ethics of Algorithms: Convening, Observation, Probability, and Timeliness', *Science, Technology, & Human Values* 41, no. 1 (January 2016): 93–117, https://doi.org/10.1177/0162243915606523; Jonathan Roberge and Louis Melançon, 'Being the King Kong of Algorithmic Culture Is a Tough Job after All: Google's Regimes of Justification and the Meanings of Glass', *Convergence* 23, no. 3 (1 June 2017): 306–24, https://doi.org/10.1177/1354856515592506.See also: Section X.X and Section Below: Algorithmic Differentiation.

[74] Verbeek, 'Materializing Morality'; Ananny, 'Toward an Ethics of Algorithms'; Roberge and Melançon, 'Being the King Kong of Algorithmic Culture Is a Tough Job after All'; Veale, Van Kleek, and Binns, 'Fairness and Accountability Design Needs for Algorithmic Support in High-Stakes Public Sector Decision-Making'; Michael Veale and Irina Brass, 'Administration by Algorithm?: Public Management Meets Public Sector Machine Learning', in *Algorithmic Regulation*, ed. Karen Yeung and Martin Lodge (Oxford University Press, 2019), 0, https://doi.org/10.1093/oso/9780198838494.003.0006.

[75] See *among others* Custers et al., *Discrimination and Privacy in the Information Society*, 352. Daniel J. Solove, 'Nothing to Hide: The False Tradeoff Between Privacy and Security (Full Text of Book)', SSRN Scholarly Paper (Rochester, NY, 31 May 2011), https://papers.ssrn.com/abstract=3976770. Maddalena Favaretto, Eva De Clercq, and Bernice Simone

the same time, if the main concern lies with the protection and safeguarding of equality as a social value, one should also look beyond concrete instances of differentiation, and consider the impact of algorithms at large in a more comprehensive manner.

ALGORITHMIC DIFFERENTIATION I – Consider a typical decision-making process, such as the hiring of a new employee by a company. In situations like these, at least two discrete instances of differentiation take place.[76] First, the company wants to hire the right candidate and to do so they want to be able to differentiate between the various persons that will apply for the position. The company will want to define those criteria and traits that an employee should possess. The totality of these traits constitutes the profile of the ideal employee, and is a first instance of differentiation, which will divide the population (at least in relation to this decision) into two categories: those people that can get hired and those that cannot. The second instance of differentiation concerns the actual application of the profile, i.e., a person who applies is either included or precluded from further screening.

At first sight, it may appear that all the elements we need to evaluate the fairness or justice of a decision-making process are already present: considering the goal in question (the hiring decision), the evaluator can assess whether the attributes that were used to distinguish between the applicants can be (reasonably) justified. Though this assessment remains a key component for the evaluation of decision-making systems – indeed, in most cases, people should be given insight in the criteria or attributes that were used to differentiate between them - the justified nature of algorithmic decision-making systems cannot simply be established by looking at this instance of differentiation alone.

ALGORITHMIC DIFFERENTIATION II – Algorithmically guided decisions are complex, multi-stage and multi-actor processes. A variety of people participate in their construction and evaluation, from data planning and data selection to system development and interpretation of the results, to the

---

Elger, 'Big Data and Discrimination: Perils, Promises and Solutions. A Systematic Review', *Journal of Big Data* 6, no. 1 (5 February 2019): 12, https://doi.org/10.1186/s40537-019-0177-4. Mann and Matzner, 'Challenging Algorithmic Profiling'; Vedder and Naudts, 'Accountability for the Use of Algorithms in a Big Data Environment'; Naudts, 'How Machine Learning Generates Unfair Inequalities and How Data Protection Instruments May Help in Mitigating Them'. Sandra Wachter, 'Affinity Profiling and Discrimination by Association in Online Behavioural Advertising', SSRN Scholarly Paper (Rochester, NY: Social Science Research Network, 15 May 2019), https://papers.ssrn.com/abstract=3388639.

[76] See also: Naudts, 'How Machine Learning Generates Unfair Inequalities and How Data Protection Instruments May Help in Mitigating Them'.

deployment, etc.[77] During every phase, choices are made that have a concrete impact on the way people are treated and judged.

The choice of data is one such instance of differentiation (e.g., what or whose data will be included or excluded from the data set?) that can have a considerable impact on a decision process. For example, the initial data set will determine what attributes will be discovered as making up our ideal applicant. The attributes a learning algorithm uncovers as being indicative of an ideal applicant may appear justified at first sight. Yet, would we still allow their use upon the discovery that those attributes were derived from a training data set that comprised of profiles of male workers only?[78] How a certain decision was reached therefore remains an important component of our assessment, and instances of differentiation that were made as part of this decision-making process may shape our assessment of whether a given decision was fair or not, justified or not. In addition, automated decision-making systems may need to be dynamic and adaptive to their environment. In such environments, the system may introduce forms of differential treatment that were not foreseen *a priori.*[79] For example, a recommender system might come to understand how a person's interests have changed over time and cater its recommendations accordingly.

By focusing solely on the attributes that comprise a profile, within a specific timeframe and with disregard of the context in which those attributes arose (prior, present, and future), other (potential) instances of *ex-ante* or *ex-post* injustice would go unnoticed.

ALGORITHMIC INEQUALITY – Instances of algorithmic differentiation interest us because they inevitably lead to a variety of (in)equalities, which can or cannot be justified.[80] They can lead to

[77] Lorena Jaume-Palasí and Matthias Spielkamp, 'Ethics and Algorithmic Processes for Decision Making and Decision Support', n.d., 19.

[78] At this point, the evaluation process can go a variety of ways. Three options appear open. First, an evaluator could consider that the decision-making process has become tainted. Regardless of how the system operates in practice, the system should not be used because the data collection process was discriminatory towards women. Second, the evaluator could condemn the decision-making process, but only insofar as the hiring process in practice would systematically disfavour women. Finally, an evaluator may agree that there was indeed an injustice during the data collection process but find that attributes found through the machine learning algorithm can still be validly used because the criteria could reasonably be considered relevant for a decision-making process. Because we have two instances of differentiation, which lead to a variety of outcomes in evaluation, it is best to consider them as highly related, but separate, nonetheless.

[79] In these cases, an audit may reveal how people are differentiated against in ways and based on attributes that were not initially foreseen by a decision-maker

[80] See also Supra, Chapter 1, Section "Algorithms and (In)equality in the Everyday."

inequality in the relationships people hold, in the opportunities people are afforded, in the (non)material goods and services people have access to, etc.[81] These (in)equalities are tied to the specific differentiation ground that is used as a decisional basis within a decision-making process. At the same time, these grounds are tied to an algorithmic group. A hiring algorithm may distinguish between decision-subjects based upon their (non-)possession of a set of characteristics valued by the hiring company. Based upon these characteristics two profiles are generated: a profile of the ideal applicant (persons with characteristics a, b and c are good hires) versus a profile of the non-ideal applicant (persons with characteristics -a, -b and -c are bad hires). Each profile represents a group of people: the totality of persons who either share or do not share the desired characteristics. Each person either classifies as an ideal or non-ideal applicant and can be assigned to one of the two groups of which they are a member. The members of the ideal group will be treated equally: they will all receive an invitation into the next round of job interviews. Members of the of the non-ideal group are also treated equally: they all get a message that their job application has unfortunately been rejected. Evidently, members of the ideal applicant group are treated differently, *read* unequally, from members of the non-ideal applicant group. When generating these inequalities, algorithms perform a restructuring function. They alter the relationships people stand in (ideal versus non-ideal applicants), and they alter the socio-economic expectations of decision-subjects (a job opportunity versus no job opportunity). These specific functions, and why they matter for the evaluation of algorithms, will be further explored in Chapter 2 of this dissertation.[82]

ALGORITHMS AS THE SUBJECT OF JUSTICE? - In his Theory of Justice, John Rawls identified as the subject of justice, the basic structure of society. He argued that the demands of justice ought to be satisfied by the major social institutions: the political constitution, as well as the principal economic and social arrangements.[83] These institutions ought to be just as they define (among others) people's "rights and duties and influence their life's prospects, what they can expect to be, and how well they can hope to do."[84] The burdens and benefits in question vary greatly, ranging from people's access to social welfare benefits, opportunities towards education, to their enjoyment of fundamental rights, etc. The basic structure of society captures a wide range of interpersonal practices, and their effects on both the socio-relational and socio-economic position people hold within society. In this sense, algorithmic decision-making systems are understood as

[81] Naudts, 'AI En Algoritmische Categorisatie: Gelijkheid En Non-Discriminatie'.

[82] See *infra* Chapter 2 Equality as an Evaluative Concept.

[83] John Rawls, *A Theory of Justice* (Cambridge, Massachusetts, London: Harvard University Press (Belknap Press), 1971), 7.

[84] Rawls, 7.

socio-technical systems whose functioning weave into the basic structure of society. Due to their ubiquity, algorithmic structures, and decision-making acts co-mediate, co-govern and co-shape society's basic institutions and are therefore subject to the demands of justice.[85]

1.2 Why Treat Algorithmically Guided Decisions Differently?[86]

People too must differentiate in the decisions they make, and when they do, their reasoning can be biased, prejudiced or discriminatory as well.[87] Moreover, like algorithms, people too can differentiate between people in unpredictable or arbitrary ways.[88] Furthermore, many algorithmic decisions differentiate upon the instructions given by their designers or developers. Why then treat algorithmic decisions differently from human ones? The latter debate comprises two distinct issues with two different solutions. First, to address the challenges algorithms raise, is an inquiry into the algorithmic nature of decision-making really needed or could regulators instead focus on the human component? This question must be answered affirmatively. There will indeed be certain elements specific to the functioning of algorithms that warrant further inspection. Second, should regulators uphold a higher level of scrutiny for algorithms? Here, the response must be negative. Both human- and machine-led decisions should meet the same standards of accountability. Every decision, whether automated or not, should be justified and reasoned for.

In this Section I want to further demonstrate how algorithmically guided decision-making systems differ from human decision-making processes, and consequently, requires policy makers to approach these systems with an altered sense of caution. Algorithms (can) be different from human decision-making in terms of a) the reasoning process involved, b) the scale at which they

---

[85] See on this point also: Iason Gabriel, 'Towards a Theory of Justice for Artificial Intelligence', *Daedalus* 151, no. 2 (forthcoming spring 2022): 3.

[86] This section is partially based on the following book Chapter: Naudts, 'AI En Algoritmische Categorisatie: Gelijkheid En Non-Discriminatie'.

[87] For a definition of these notions, see Chapter 2, Section 2.3 "Algorithmic Inequality: Related Notions."

[88] See for instance the much-used example of judges allegedly issuing rougher punishment before their breaks as they tire and get hungry. Zoë Corbyn, 'Hungry Judges Dispense Rough Justice', *Nature*, 11 April 2011, https://doi.org/10.1038/news.2011.227. In reference to: Shai Danziger, Jonathan Levav, and Liora Avnaim-Pesso, 'Extraneous Factors in Judicial Decisions', *Proceedings of the National Academy of Sciences* 108, no. 17 (26 April 2011): 6889–92, https://doi.org/10.1073/pnas.1018033108.

operate, c) the complexity of information they consider and d) the ways through which they (could potentially) interconnect with one another.[89]

1.2.1 Differences in Reasoning

The information algorithms and humans take aboard as being relevant can differ. For certain algorithmically guided decision-making processes, the knowledge basis they use to differentiate between people can be generated by data-driven knowledge discovery techniques.[90] For example, analytics can be deployed to uncover hidden patterns and correlations within large data sets. The relationships these technologies discover may have never been discovered or thought of by people otherwise. Machines simply have more time, processing power, and (intellectual) capacity to sift through, and explore substantial amounts of data, than humans. For instance, online lending company ZestFinance discovered through learning techniques that people who fill in their loan applications using capital letters only were more likely to default.[91] In this context, Wachter makes a further distinction between two types of differentiation grounds.[92] On the one hand there are differentiation grounds defined by characteristics and definition rules that are humanly comprehensible. Examples of such categories are "married sophisticates" or people living in poverty.[93] On the other hand, there are differentiation grounds defined by characteristics and

---

[89] The argument could be made that the properties described below are not exclusive to algorithms. For example, decisions-makers could produce highly complex decision-making systems through analogue means. Likewise, certain highly complex automated decision-making systems could be reproduced via non-automated means given enough time or resources. Still, algorithmic techniques have allowed decision-makers to incorporate these properties to their advantage with greater ease. Hence, the negative social impact I link to these properties is more likely to occur through, and because of, technology.

[90] Not every algorithmically guided decision works like this. Sometimes the algorithm simply performs a knowledge application function, rather than a knowledge discovery one.

[91] Stephen F. DeAngelis Solutions Enterra, 'Artificial Intelligence: How Algorithms Make Systems Smart', *Wired*, 5 September 2014, https://www.wired.com/insights/2014/09/artificial-intelligence-algorithms-2/; Deborah Gage, 'Big Data Uncovers Some Weird Correlations', *Wall Street Journal*, 23 March 2014, sec. Special, https://online.wsj.com/article/SB10001424052702303369904579423132072969654.html; Steve Lohr, 'Banking Start-Ups Adopt New Tools for Lending', *The New York Times*, 19 January 2015, sec. Technology, https://www.nytimes.com/2015/01/19/technology/banking-start-ups-adopt-new-tools-for-lending.html. Solutions, 'Artificial Intelligence'. In response, ZestFinance Chief Executive Douglas Merril explained that it was up to data scientists to verify whether or not their findings made sense. For more examples of in the credit sector: Hurley and Adebayo, 'Credit Scoring in the Era of Big Data'.

[92] Wachter, 'The Theory of Artificial Immutability', 6. See also Chapter 2, Section: Algorithmic Equality of Whom?

[93] Data brokers combine and classify consumers in novel social segments based upon their interests. For instance, the FTC discovered the use of the group "Married Sophisticates." This group included thirty-something couples in the

definition rules that are not humanly comprehensible or cannot be meaningfully interpreted by humans due to their scale, volume, complexity, or source. Examples of this kind are differentiating acts based upon browsing behaviour, or other electronically monitored signals.[94] The capacity for algorithmic techniques to uncover new differentiation grounds hitherto unknown by humans has been identified by scholars and policy makers as one of the core challenges the algorithmic environment poses to equality (and non-discrimination).[95]

To some extent, this capacity of data-driven analytics techniques to investigate and make use of massive amounts of data in many different ways, represents a way of reasoning that can be different, and sometimes counterintuitive, to humans.[96] Furthermore, and as aforementioned, learning techniques discover correlations, rather than causations, which will serve as an adequate basis for action.[97] The validity of inductive and correlation-based reasoning has been questioned

---

upper-middle class with no children. Edith Ramirez et al., 'Data Brokers: A Call for Transparency and Accountability' (Federal Trade Commission, 2014), 47. Wachter, 'The Theory of Artificial Immutability', 6.

[94] Wachter, 'The Theory of Artificial Immutability', 6. In reference to: Jenna Burrell, 'How the Machine 'thinks': Understanding Opacity in Machine Learning Algorithms', *Big Data & Society* 3, no. 1 (2016): 2053951715622512, https://doi.org/10.1177/2053951715622512; Tim Miller, 'Explanation in Artificial Intelligence: Insights from the Social Sciences', *ArXiv:1706.07269 [Cs]*, 22 June 2017, http://arxiv.org/abs/1706.07269; Brent Mittelstadt, Chris Russell, and Sandra Wachter, 'Explaining Explanations in AI', in *Proceedings of the Conference on Fairness, Accountability, and Transparency*, FAT* '19 (New York, NY, USA: Association for Computing Machinery, 2019), 279–88, https://doi.org/10.1145/3287560.3287574; Christoph Molnar, *Interpretable Machine Learning*, accessed 8 September 2022, https://christophm.github.io/interpretable-ml-book/.

[95] See for example: Laurens Naudts, 'Fair or Unfair Differentiation? Luck Egalitarianism as a Lens for Evaluating Algorithmic Decision-Making.' (Data for Policy Conference, London, 2017); Naudts, 'How Machine Learning Generates Unfair Inequalities and How Data Protection Instruments May Help in Mitigating Them'; Vedder and Naudts, 'Accountability for the Use of Algorithms in a Big Data Environment'; Naudts, 'AI En Algoritmische Categorisatie: Gelijkheid En Non-Discriminatie'; Wachter, 'Affinity Profiling and Discrimination by Association in Online Behavioral Advertising'; Leese, 'The New Profiling: Algorithms, Black Boxes, and the Failure of Anti-Discriminatory Safeguards in the European Union'; Mann and Matzner, 'Challenging Algorithmic Profiling'; Borgesius, 'Strengthening Legal Protection against Discrimination by Algorithms and Artificial Intelligence'; Gerards and Zuiderveen Borgesius, 'Protected Grounds and the System of Non-Discrimination Law in the Context of Algorithmic Decision-Making and Artificial Intelligence', 2 November 2020. On a policy level, see: Ad Hoc Committee on Artificial Intelligence (CAHAI), 'Feasibility Study'.

[96] See also: Cade Metz, 'In Major AI Breakthrough, Google System Secretly Beats Top Player at the Ancient Game of Go', *Wired*, accessed 19 May 2022, https://www.wired.com/2016/01/in-a-huge-breakthrough-googles-ai-beats-a-top-player-at-the-game-of-go/.

[97] Mireille Hildebrandt and Bert-Jaap Koops, 'The Challenges of Ambient Law and Legal Protection in the Profiling Era: The Challenges of Ambient Law and Legal Protection in the Profiling Era', *Modern Law Review* 73, no. 3 (May 2010): 432, https://doi.org/10.1111/j.1468-2230.2010.00806.x; Mittelstadt et al., 'The Ethics of Algorithms: Mapping the Debate', 5.

on a scientific level.[98] Advances in the development of machine learning techniques based on principles of causal learning and inference hope to further strengthen the reliability and robustness of decision-making.[99] Similarly, human reasoning can be used to complement data-driven information in order to establish a stronger knowledge base. Certain grounds used to differentiate people may be more robust than others. Yet, even in the case a feature appears sufficiently robust, or even approximates a causal connection, it does not follow that its use as a differentiation ground is legitimate.[100]

1.2.2 The Complexity of Decision-Making Grounds

Next to enabling human decision-makers to take into account new knowledge, algorithmic techniques - aided by increased computing power – can also process information and decisions that are far more complex in nature. For instance, a learning algorithm may discover that a multitude of features are relevant for a particular decision-making process. While a human decision-maker may have difficulty taking all these features into account, automated techniques do not. The administrative overload caused by complex decision-making processes in analogue settings is far less an issue within the automated environment.

Algorithms also allow complex and human-generated profiles to be applied onto individuals with greater ease. For example, the COMPAS recidivism tool, which saw use in the United States, allocated risk scores to defendants in the context of criminal justice procedures.[101] Those scores

---

[98] Mittelstadt et al., 'The Ethics of Algorithms: Mapping the Debate', 5.

[99] 'Causality and Machine Learning', *Microsoft Research* (blog), accessed 19 May 2022, https://www.microsoft.com/en-us/research/group/causal-inference/.

[100] On this point, I would like to thank Dr. Reuben Binns for raising this challenge in our conversations. In particular, if we accept the use of some correlations as a basis for decision-making, and a correlation appears as robust across multiple models - and upon the condition that it is not associated with an existing category of discrimination or oppression - would it not be a legitimate ground to use? Throughout this dissertation, my argument will be that one should not only look at the attributes, or their reliability, in order to assess a decision-making system. One should also consider the likely harms a system can generate for decision-subjects. If these harms are too great, then one should perhaps reconsider the deployment of the system to begin with. This can also be said in case true causal links were found.

[101] See among others: Mattu et al., 'How We Analyzed the COMPAS Recidivism Algorithm'; Northpointe, 'COMPAS Risk & Need Assessment System' (Northpointe, 2012), http://www.northpointeinc.com/files/downloads/FAQ_Document.pdf; Equivant, 'Practicioner's Guide to COMPAS Core' (Equivant (Northpointe Inc.), 4 April 2019), https://www.equivant.com/wp-content/uploads/Practitioners-Guide-to-COMPAS-Core-040419.pdf; Store, 'COMPAS Recidivism Risk Score Data and Analysis'; Julia Angwin et al.,

were the result of a complex equation following a questionnaire filled in by the defendants. That questionnaire comprised 137 questions and was developed based upon research in behavioural and social science. Still, automated techniques enabled a more efficient analysis of the answers provided to these questions. Case workers no longer had the administrative burden to examine, weigh and score the answers provided. So even if the profiles themselves remain entirely human made, there is more freedom in making these profiles more intricate when their analysis and application can be outsourced to computer and machines.[102]

1.2.3 Group-level Knowledge

Data-driven knowledge is often derived from group- or population-level insights.[103] Drawing from single or multiple sources of information, an aggregate of data is accumulated, whereby the aggregate of data is further explored to discover new knowledge. For instance, "individuals are clustered [differentiated] into meaningful groups (or classes) according to their behaviour, preferences, and other characteristics."[104] The discovered knowledge (which can be seen as the totality of extracted information) in turn, can be applied back onto either individual persons or a group of persons.[105]

Group-level knowledge often takes the form of approximations and generalisations. Such generalisations can be beneficial because they can reduce information overload.[106] Yet, in doing so, they function as a predictor of behaviour (or other characteristics) at the level of the group and

'Machine Bias', ProPublica, 23 May 2016, https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing?token=lZ_nPrh6oVJEnMzcTH1Jr59Ibe3K8XZC.

[101] See for example: Salomé Viljoen, 'A Relational Theory of Data Governance', *The Yale Law Journal*, 2021, 578. See also: Ananny, 'Toward an Ethics of Algorithms'; Mittelstadt et al., 'The Ethics of Algorithms: Mapping the Debate', 5. Hildebrandt, 'Defining Profiling'.

[102] Of course, where a decision becomes more complex, this might also translate to a heightened level of complexity on a technical level. See *supra* Section 1.2.1 Black Boxes: An Obstacle Towards Evaluation?

[103] See for example: Viljoen, 'A Relational Theory of Data Governance', 578. See also: Ananny, 'Toward an Ethics of Algorithms'; Mittelstadt et al., 'The Ethics of Algorithms: Mapping the Debate', 5. Hildebrandt, 'Defining Profiling'.

[104] Mittelstadt, 'From Individual to Group Privacy in Big Data Analytics', 476.

[105] It should be noted that the initial source of information collection should not necessarily be related to persons. Data of inanimate objects, such as traffic light data, or natural phenomena, such as climate change, could equally allow conclusions to be made that could affect decisions regarding people. See also: Hildebrandt, 'Defining Profiling'.

[106] Hildebrandt, 30.

not at the level of the individual.[107] Put differently, they provide insight into the *likely* relations people hold with others, rather than individual-specific insights.[108] In the data-driven environment, these (group-level) generalisations are often referred to as profiles, and the practices that lead to their creation and application as profiling.[109]

DISTRIBUTIVE VERSUS NON-DISTRIBUTIVE GENERALISATIONS - Group-level generalisations can be distributive or non-distributive.[110] On the one hand, distributive generalisations assign properties to a group that are actually and unconditionally manifested or shared by all members of the group.[111] Non-distributive generalisations on the other hand are framed in terms of "probabilities and averages and medians, or significant deviances from other groups" and "based on comparisons of members of the group with each other and/or on comparisons of one particular group with other groups."[112] The properties of non-distributive profiles apply to individuals as members of the reference group, and not because they actually exhibit these properties in reality.[113] An insurance company might categorise individual A with characteristics X and Y as being unhealthy, because at the group level it was determined that individuals who exhibit characteristics X and Y are unhealthy. The individual might however also exhibit a characteristic Z which would negate the presence of X and Y, and instead indicate that they are healthy.[114]

---

[107] Tal Zarsky, 'Transparent Predictions', SSRN Scholarly Paper (Rochester, NY: Social Science Research Network, 10 September 2013), https://papers.ssrn.com/abstract=2324240; Mittelstadt, 'From Individual to Group Privacy in Big Data Analytics'. The algorithm can moreover change the composition of the group on the basis of new data, or as a result of improvements in terms of accuracy.

[108] Viljoen, 'A Relational Theory of Data Governance', 578.

[109] The notion *profiling* may be used to denote the process of analysing a given subject or group to generate a profile, as well as the application of said profile. For more information on the term, reference can be made to: Hildebrandt and Gutwirth, *Profiling the European Citizen*. Hildebrandt, 'Defining Profiling'. And the response of: David-Olivier Jaquet-Chiffelle, 'Direct and Indirect Profiling in the Light of Virtual Persons', in *Profiling the European Citizen: Cross-Disciplinary Perspectives*, ed. Mireille Hildebrandt and Serge Gutwirth, 2008, 34–43.

[110] Vedder, 'KDD', 277. ANTON VEDDER, 'WHY DATA PROTECTION AND TRANSPARENCY ARE NOT ENOUGH WHEN FACING SOCIAL PROBLEMS OF MACHINE LEARNING IN A BIG DATA CONTEXT', in *BEING PROFILED*, ed. EMRE BAYAMLIOĞLU et al., COGITAS ERGO SUM: 10 Years of Profiling the European Citizen (Amsterdam University Press, 2018), 42–45, https://doi.org/10.2307/j.ctvhrd092.10.See also on this distinction: Patrick Allo, 'The Epistemology of Non-Distributive Profiles', *Philosophy & Technology* 33, no. 3 (1 September 2020): 379–409, https://doi.org/10.1007/s13347-019-00360-z.

[111] Vedder, 'KDD', 277.

[112] Vedder, 277.

[113] Vedder, 277.

[114] See also Chapter 6, Section 6.2.2. "The Problematic Nature of Generalisations: Deindividualisation and Misrepresentation."

OVER AND UNDER-INCLUSIVE – Generalisations can moreover be over- or under-inclusive. A generalisation is over-inclusive when it captures more people than necessary or intended. In our example, the main goal of the insurance company is to distinguish between people that are healthy and those that are unhealthy. The unhealthy category captures individuals in the possession of X and Y. Yet because no attention was paid to Z in defining that category, it also captures healthy people who do not belong to the club of unhealthy people (over-inclusive). Likewise, there might be a person who exhibits characteristic W, which is an indicator of a highly unhealthy lifestyle, but not characteristics X and Y. Whereas that person should have ideally been included as unhealthy, they are nonetheless considered healthy (under-inclusive). In short, profiles tend to be non-universal.[115]

PERSONALISED DECISIONS? - Most decision-making systems applied to individuals are unlikely to have captured all that is available about those individuals in reaching their decision. Moreover, in many decision-making instances, the decision-subject is unable to indicate their specific wants, needs, preferences or expectations directly to a decision-maker. Instead, the decision-maker will try to individualise or personalise a specific decision by relying on group-level information concerning others of whom the decision-maker assumes they are similar to the decision-subject. As individualised as the notion *personalised* decision-making therefore sounds, the practice still (or currently) concerns a more specified application of a generalisation. Considering generalisations are phrased in ways that they do not capture merely one individual, but a multitude of individuals, personalised information remains group-level information.[116]

[115] See also: Frederick Schauer, *Profiles, Probabilities and Stereotypes* (Cambridge, Massachusetts; London: The Belknap Press of Harvard University Press, 2003).

[116] The notion personalisation might still be useful to designate the goal of a decision-maker to cater their services as closely as possible to the needs of their individual end-users.

1.2.4 Scale and Amplification

A singular decision can be considered unfair or unjust for a variety of reasons. A decision may have been based on discriminatory attributes, or a person or group may have been judged on the basis of information that was inaccurate, irrelevant, or seemingly arbitrary.

Technologies can be rolled out at a large scale effortlessly. The impact of algorithms, and the injustice they introduce, might therefore be more swiftly and easily amplified. Consider two instances of decision-making, one is automated, and the other is not. Suppose the president of a prison parole board is racist. Whenever a person of colour is placed before the board, that person (unknowingly) starts at a disadvantage. Even in the case they can show the same record of good behaviour as white defendants, they are less likely to be given parole. Suppose now that a software company has developed an algorithm that will guide parole boards in evaluating the recidivism risk of defendants. After a satisfactory trial period in one prison, the system is rolled out over the entire country. Unbeknownst to the prison board and the system's developers, the algorithm is biased towards people of colour too. In both cases, the rejection of a person's request for parole on ethnic grounds is a moral wrong.[117] Yet, in the scenario of automation, the unfair differential treatment experienced by people of colour is amplified. Because the moral wrong is no longer tied to one single individual, but to every entity who uses the software, its immediate impact is more structural and systemic in nature.[118]

In addition, the data that drive these processes, including secondary information derived through algorithmic knowledge discovery techniques, can be bought, sold, and transferred.[119] Within such an interconnected environment, the function or value that a particular piece of information holds can fluctuate. For example, a specific piece of information concerning a person or group is likely to hold more value where it has great predictive power over a range of decisions. In a similar vein, due to the interconnected nature of certain algorithmically guided decision-making processes, or the interconnected nature of the sectors in which these decisions are made, one piece of (algorithmically generated) information could trigger a chain reaction.

---

[117] Caution is in order however: I do not want to promote pure consequential reasoning where harm is defined solely in relation to the scale of its consequences.

[118] Two points of attention. First, it should be noted that racism is already a structural and institutional form of oppression. Second, and admittedly, persons or entities in a position of great power, such as a government or dictator, also have the ability to issue large-scale decisions with an immediate, structural negative impact.

[119] See for instance: Ramirez et al., 'Data Brokers: A Call for Transparency and Accountability'.

THE NEGATIVE SPIRAL - In this context, the "negative spiral" argument, introduced by Citron and Pasquale, considers the possibility for algorithmic decision-making processes to generate great disparity between people's life path in a way that does not correspond to their actual differences.[120] Suppose the AlgoBank financial institution allocates a score to individuals A and B to determine whether or not they can be granted a loan. Two people with similar credentials apply for a loan at the AlgoBank. Both persons would be capable of repaying the loan they apply for. The learning algorithm the AlgoBank employs however attaches great significance to a person's pay grade. Person A earns a hundred euro more than Person B. Person A is lucky and falls within the AlgoBank's good graces: the algorithms classify A as a trustworthy applicant. Person B is less lucky though and needs to seek their fortunes elsewhere. The negative spiral does not stop at this point. The reach of AlgoBank's algorithmic decision goes further. Health insurance company AlgoSecure buys data from the AlgoBank. Because Person B has not received a loan, they are now seen as presenting higher health risk. The AlgoSecure algorithm sees a strong correlation between creditworthiness, poverty, and health. That being the case, Person B must pay a higher premium. Unfortunately, B cannot pay the premium and they are left without health insurance. Such misfortune does not befall A, who can easily get the insurance. Not long after their respective encounter with the insurance company, both B and A fall ill. Having no insurance and no money for expensive medicine, B develops a chronic lung disease and they are no longer able to perform their job. Person A develops the same illness, gets treated quickly and soon recovers. Despite the fact that differences in their initial starting position were minute, one negative algorithmic decision had a snowball effect. Person B faces far more socio-economic hardship than person A. Due to their weaker socio-economic position, B feels they are shunned within society, while their chronical illness further contributes to their hardship. Taken as a whole, the negative impact experienced by one party is so disproportionate that it cannot be justified in reference to those initial differences.[121]

The negative spiral argument illustrates how a given trait can have an increased and amplified impact. Person B not only faces greater socio-economic inequality as time progresses. Feeling shunned, they also experience social inequality in the relations they hold.[122] It might not suffice

---

[120] Danielle Keats Citron and Frank Pasquale, 'The Scored Society: Due Process for Automated Predictions', *Washington Law Review* 89, no. 1 (2014): 33. Tal Z. Zarsky, 'Understanding Discrimination in the Scored Society', *Wash. L. Rev.* 89, no. 4 (2014): 1405.

[121] Citron and Pasquale, 'The Scored Society: Due Process for Automated Predictions', 2014; Zarsky, 'Understanding Discrimination in the Scored Society', 1405.

[122] Citron and Pasquale, 'The Scored Society: Due Process for Automated Predictions', 2014; Zarsky, 'Understanding Discrimination in the Scored Society'. There are limits to this interconnected nature of the digital environment however.

for policy makers to focus on one specific decision-making instance. Even if every singular instance of differentiation within the aforementioned decision-chain could be justified, the question must be raised whether the final outcome of that entire chain can find justification too. The negative spiral also confronts policy makers with a problem of uncertainty vis-à-vis the pervasive and successive effects algorithms can have over time. The long-term effects of automated decision-making are difficult to predict and address *a priori*. Instead, policy makers should consider the introduction of periodic monitoring and course-correction mechanisms.[123]

Due to their amplifying, interconnected and group-level nature, the inegalitarian impact algorithms produce, and the social change these inequalities initiate, should not be underestimated.[124] Algorithms have already been shown to reinforce historical injustice. At the same time, these technologies have expanded decision-makers' toolset. They can now discover seemingly random groups as relevant for specific decision-goals or, they can increase the complexity of the decisional grounds they rely upon. The inequalities imposed onto seemingly random and non-socially salient groups may alter the structure of society alongside new and currently undefined social strata with unforeseen and (potentially) unjustifiable consequences as a result.[125] Furthermore, due to the opacity and complexity of algorithmic systems, processes of social transformation may take place in ways that are more discreet and imperceptible and only visible over time.[126]

1.2.5 The Black Box: An Obstacle Towards Evaluation?

While most algorithmically guided decision-making processes rely upon differentiation, they differ in their explicability. Certain algorithmic decision-making systems have been developed as

---

Algorithmic systems are contextually dependent. For instance, a hiring system specifically built for the Belgian market, could not be placed in the Netherlands without making the necessary changes. Because the system was developed with the Belgian hiring culture in mind, it would consider factors that companies in The Netherlands may not find important.

[123] It is true that individuals might end up in a negative spiral without automation. Still, in an environment where data functions as an easily exchangeable currency in online transactions, decisions can be more easily interlinked, and hence the impact of one piece of information is more likely to be exchanged and adopted by others. Furthermore, due to the opacity of the algorithmic environment, inequalities may be more difficult to detect and harder to trace back their exact point of origin.

[124] Mittelstadt et al., 'The Ethics of Algorithms: Mapping the Debate', 5.

[125] Vedder and Naudts, 'Accountability for the Use of Algorithms in a Big Data Environment', 210.

[126] Lydia T. Liu et al., 'Delayed Impact of Fair Machine Learning', *ArXiv:1803.04383 [Cs, Stat]*, 12 March 2018, http://arxiv.org/abs/1803.04383.

a black box, hiding their internal logic from the user and sometimes even the designer.[127] In these black-box situations, decision-makers are unable to provide a level of explanation or justification one would typically expect under normal standards of accountability and good governance. In this regard, the use of black-box systems can be criticised and the lack of explainability associated with them can even be seen as form of procedural injustice or unfairness.[128] Yet, as it stands, these systems are used in practice and in critical sectors.[129]

In the face of missing information, it is indeed less evident to make a proper assessment of decision-making procedures. Ideally, some degree of transparency is needed to evaluate a decision-making procedure, and the more information we have, the more sound our evaluation can be.[130] Even if we do not have access to all information, the black-box nature of algorithms does not render every evaluative exercise obsolete or impossible.

First, the argument could be made that the internal logic, or at least an approximation thereof, of complex decision-making systems should be made available as a precondition for their use.[131] Once sufficient insight is given into the inner workings of a system, an ex-ante and ex-post evaluation of their functioning can be made. Second, in the case algorithmic systems produce negative consequences, one might expect that their injustice will show itself over time. While such

---

[127] Riccardo Guidotti et al., 'A Survey of Methods for Explaining Black Box Models', *ACM Comput. Surv.* 51, no. 5 (August 2018): 93:1-93:42, https://doi.org/10.1145/3236009. The black box problem is particularly present in unsupervised and deep learning algorithms.

[128] High-Level Expert Group on Artificial Intelligence, 'Ethics Guidelines for Trustworthy AI'.

[129] Frank Pasquale, *The Black Box Society: The Secret Algorithms That Control Money and Information*, First Harvard University Press paperback edition (Cambridge, Massachusetts London, England: Harvard University Press, 2015); Katherine Fink, 'Opening the Government's Black Boxes: Freedom of Information and Algorithmic Accountability', *Information, Communication & Society* 21, no. 10 (3 October 2018): 1453–71, https://doi.org/10.1080/1369118X.2017.1330418; Guidotti et al., 'A Survey of Methods for Explaining Black Box Models'; Juan Manuel Durán and Karin Rolanda Jongsma, 'Who Is Afraid of Black Box Algorithms? On the Epistemological and Ethical Basis of Trust in Medical AI', *Journal of Medical Ethics* 47, no. 5 (1 May 2021): 329–35, https://doi.org/10.1136/medethics-2020-106820.

[130] The forum of transparency need not be decision subjects. They can also be audit organisations or external review boards.

[131] Of course, people too might reason in a black-box manner. Still, we can encourage persons via mechanisms of accountability to make their reasoning process and underlying assumptions and motivations explicit. This could give us an opportunity to scrutinise whether the outcome of the decision followed a reasoning process that was consistent and coherent and in adherence to the governing norms of society. On human bias in decision-making, see also: John Zhuang Liu and Xueyao Li, 'Legal Techniques for Rationalizing Biased Judicial Decisions: Evidence from Experiments with Real Judges', *Journal of Empirical Legal Studies* 16, no. 3 (2019): 630–70, https://doi.org/10.1111/jels.12229.

processes might unfortunately go unnoticed for a long time, as soon as their impact does become clearer, the decision-making process can be made subject to a concrete evaluation. In certain cases, the presence of an undesired outcome may already be sufficient to prohibit the further use of a specific system. For instance, a soon as we discover that a decision-making procedure produces biased results against people with a certain ethnic background, we could question its functioning and deployment.[132] This demonstrates how algorithmic decision-making systems should not be evaluated by solely considering their logic, inner workings or differentiating features. For example, one might more easily gain insight into the data that were used within a system, rather than the operations performed on said data. Having access to the underlying data may therefore already bring certain injustices to the forefront. Likewise, given the context and domains in which an algorithmic system will be used, we can evaluate the desirability of their deployment by looking at the harms those systems are likely to impose onto decision-subjects.

Of course, certain challenges will undoubtedly remain. Highly adaptive systems may be difficult to evaluate at a particular moment in time. Their impact may need to be monitored and reviewed over extended periods of time and if necessary, course-corrected to ensure they continue to correspond to our sense of justice.[133] At the same time, our sense of justice may change as well. Systems we thought were just may appear unjust in the future. We should thus remain critical towards our own capacity to evaluate systems.

1.3 Algorithmic Power: Infrastructures, Decisional Control and Knowledge Creation.

The notion of power holds significant importance for the evaluation of algorithmically guided decision-making processes. Actors in a position of power typically hold (some) capacity to govern and decide over persons placed in a position of lesser power. The former may therefore exert influence over the (in)equalities carried by the latter. Hence, it is worth investigating the positions of power that characterise the algorithmic environment.

---

[132] Of course, it would be interesting to gain insight into the system to uncover why these biased results were produced in the first place, and what measures we can take to avoid them in the future. Similarly, locating where a system went wrong can help us when we need to allocate responsibility for the harm caused.

[133] For example, in order to verify whether or not the impact they have on people is negative, or whether this impact is primarily felt by one particular group, and hence might obtain a structural and systemic character. See also: Liu et al., 'Delayed Impact of Fair Machine Learning'.

Not everyone has the capacity to automate decision-making procedures. To do so, one needs access to the means of algorithmic knowledge production or application, such as data, hardware, analytics software, etc. These means of knowledge discovery and production are a source of power. Furthermore, when control over these means is centralised, the power to restructure society becomes centralised too. In these instances, the social impact of the choices made by those in control can be even more significant. Three distinct, though highly related, and often overlapping positions of power stand out: infrastructural control, decisional control, and control over knowledge.[134]

1.3.1 Infrastructural Control

Infrastructural control refers to an entity's access to the resources required to implement automated decision-making processes. The most evident resource is data. Fisher and Streinz argue that the digital environment is characterized by unequal control over data: "By unequal control over data, we not only mean having or not having data, but also having or not having power over deciding what becomes and what does not become data."[135] Fisher and Streinz refer to this as the power entities hold to datafy, which relates to the unequal control entities hold over the "infrastructures that generate, shape, process, store, transfer, and use data."[136] Access to inferential analytics has become a prerequisite for harvesting the knowledge hidden within large datasets. The infrastructure necessary to implement automation pertains more than data and software alone. It also entails having access to hardware, as well as human resources.[137]

Many different actors operate within the digital space, and they can form partnerships in order to combine the respective control they hold over the various means of data and knowledge extraction and production. For example, not every government will have the internal ICT resources needed to make a move towards automation by developing their own tools. As an alternative, they may outsource certain administrative tasks to third parties, or buy eGovernment tools from third-party vendors. Here, hybrid structures of power manifest themselves. The

---

[134] While the power dynamics described are important, power arises within larger structures, through the actions and support of many. Iris Marion Young, *Justice and the Politics of Difference* (Princeton, New Yersey: Princeton University Press, 1990), 25–31.

[135] Angelina Fisher and Thomas Streinz, 'Confronting Data Inequality', SSRN Scholarly Paper (Rochester, NY: Social Science Research Network, 1 April 2021), https://doi.org/10.2139/ssrn.3825724.

[136] Fisher and Streinz.

[137] See also: Kate Crawford and Vladan Joler, 'Anatomy of an AI System', Anatomy of an AI System, accessed 26 January 2022, http://www.anatomyof.ai.

infrastructural power required to exercise control upon people's beings and doings is no longer reserved to governments alone. To use the wording of Crawford and Joler, "global mega-companies", such as Facebook, Google or Amazon have sufficient internal means to maintain a large sphere of influence without the need to engage in external partnerships. They have been able to expand their activities and reach, including their position as data collector and provider.[138] The choices made by those having infrastructural control can have a profound impact on the entire decision-making chain, and this is perhaps most true for those who hold power over the means of data collection, production, and extraction. Even in those cases where, at first sight, the decisional responsibility over the deployment of an automated process lies in the hands of a third-party, the choices of those who hold power over the means of data production and extraction, such as the instructions they give to data annotators, or the values they integrate during the design of a system, will have a considerable impact on a decision's outcome.[139] The power to decide whose data will or won't be collected is another excellent case in point.[140] D'Ignazio and Klein, in reference to feminist geographer Joni Seager, rightfully reemphasise the importance of the adage

[138] See for example: Carole Cadwalladr and Emma Graham-Harrison, 'How Cambridge Analytica Turned Facebook "Likes" into a Lucrative Political Tool', *The Guardian*, 17 March 2018, sec. Technology, https://www.theguardian.com/technology/2018/mar/17/facebook-cambridge-analytica-kogan-data-algorithm; Stephanie Kirchgaessner, 'Cambridge Analytica Used Data from Facebook and Politico to Help Trump', *The Guardian*, 26 October 2017, sec. Technology, https://www.theguardian.com/technology/2017/oct/26/cambridge-analytica-used-data-from-facebook-and-politico-to-help-trump; Georg Riekeles, 'I Saw First-Hand How US Tech Giants Seduced the EU – and Undermined Democracy', *The Guardian*, 28 June 2022, sec. Opinion, https://www.theguardian.com/commentisfree/2022/jun/28/i-saw-first-hand-tech-giants-seduced-eu-google-meta; Glenn Greenwald et al., 'Microsoft Handed the NSA Access to Encrypted Messages', *The Guardian*, 12 July 2013, sec. US news, https://www.theguardian.com/world/2013/jul/11/microsoft-nsa-collaboration-user-data; Ramirez et al., 'Data Brokers: A Call for Transparency and Accountability'; Jack Nicas, Raymond Zhong, and Daisuke Wakabayashi, 'Censorship, Surveillance and Profits: A Hard Bargain for Apple in China', *The New York Times*, 17 May 2021, sec. Technology, https://www.nytimes.com/2021/05/17/technology/apple-china-censorship-data.html.

[139] Milagros Miceli, Martin Schuessler, and Tianling Yang, 'Between Subjectivity and Imposition: Power Dynamics in Data Annotation for Computer Vision', *Proceedings of the ACM on Human-Computer Interaction* 4, no. CSCW2 (14 October 2020): 1–25, https://doi.org/10.1145/3415186. As Denton and others observe: *"countless subjective and value-laden decisions go into the construction of a dataset. Yet, once a dataset is released and becomes established enough to seamlessly support research and development, the contingent conditions of creation tend to be lost or taken for granted. Once naturalized, datasets are more likely to be treated as neutral or scientific objects and uncritically adopted within daily work routines."* Emily Denton et al., 'Bringing the People Back In: Contesting Benchmark Machine Learning Datasets', *ArXiv:2007.07399 [Cs]*, 14 July 2020, 3, http://arxiv.org/abs/2007.07399.

[140] See also Chapter 6, Section 6.1.5. Algorithmic Oppression.

"What gets counted counts" in the digital world.[141] The basis for decision-making is that was has been counted or collected, and that what does not get counted or collected remains invisible.[142]

1.3.2 Decisional Control

Having access to the resources to digitise or automate decision-making processes is one thing. Control over how those resources are used, is another. Decisional control refers to an entity's influence over the design and deployment of an automated decision-making process. Before automation can take place, a whole host of questions will need to be addressed: What decisions do we want to automate, and for which purposes? Who will be subjugated to the decision? Will there be human oversight? This process may also include questions that are related to the infrastructures required for those ends, including what data are needed or whose data should (or should not) be collected? Decision-making processes will differ greatly depending on how these questions are answered, and decision-makers may often be unaware over the impact or value their choices hold, such as taking the subjective assumptions of oneself and others as an objective or self-evident truth. [143]

---

[141] Catherine D'Ignazio and Lauren F. Klein, *Data Feminism*, Strong Ideas Series (Cambridge, Massachusetts: The MIT Press, 2020), chap. 4, https://mitpress.mit.edu/books/data-feminism.

[142] D'Ignazio and Klein, *Data Feminism*. See also: Caroline Criado Perez, 'The Deadly Truth about a World Built for Men – from Stab Vests to Car Crashes', *The Guardian*, 23 February 2019, sec. Life and style, https://www.theguardian.com/lifeandstyle/2019/feb/23/truth-world-built-for-men-car-crashes. Caroline Criado Perez, *Invisible Women: Exposing Data Bias in a World Designed for Men*, 01 ed. (London: Chatto & Windus, 2019). One should not read this statement as an argument in favour of more data collection. Certain groups want to remain invisible. Moreover, more data collection would come into conflict with other values, such as privacy and data protection, and the principle of data minimisation. What is more important is to have *smart* data collection, and accompanying measures that can account for the interests of those that could not be counted. See also Chapter 6, Section 6.1.5. Algorithmic Oppression.

[143] Certain decisions may present themselves as value-laden more evidently than others. A facial recognition system built to detect "beauty" for instance may not be immediately seen or approached as objective. Yet, certain classification mechanisms remain uncontested despite obviously being value-laden, such as the classification of people alongside socially-constructed categories, such as race and gender, which are typically presented as undisputed, while they are not. Milagros Miceli et al., 'Documenting Computer Vision Datasets: An Invitation to Reflexive Data Practices', in *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency* (FAccT '21: 2021 ACM Conference on Fairness, Accountability, and Transparency, Virtual Event Canada: ACM, 2021), 169, https://doi.org/10.1145/3442188.3445880. In reference to: Morgan Klaus Scheuerman et al., 'How We've Taught Algorithms to See Identity: Constructing Race and Gender in Image Databases for Facial Analysis', *Proc. ACM Hum.-Comput. Interact.* 4, no. CSCW1 (2020), https://doi.org/10.1145/3392866. Crawford too highlights the problem of how "gender, race, and sexuality are falsely assumed to be natural, fixed, and detectable biological categories." Kate

Those with infrastructural control will often have some degree of decisional control, whether directly or indirectly. A political party might decide they want to run a targeted advertising campaign, but their choices when doing so will rely upon the options provided to them by the platform on which they will run their campaign. At the same time, an actor with infrastructural control might make available their data but remain unaware as to the exact purposes for which those data will be used. This can make algorithmic systems complex hybrid structures, where many hands exercise a degree of control over how a decision will be governed.[144]

It also important to consider the type of decision a given entity holds control over. For example, what are the goods and services a decision-maker reigns over, and what is their social sphere of influence? Public actors for instance hold considerable power over their citizens in relation to decisions that concern the structuring of the welfare state, whereas private actors may have control over other important domains of our daily lives, such as loans or job opportunities.

1.3.3 Control over Knowledge

Finally, actors differ in the control they have over knowledge, which can take at least two shapes. The first asymmetry related to knowledge concerns the power to understand and explain decision-making systems. Decision-making systems are hallmarked by great complexity and often opacity.[145] From a technical perspective, these systems are accessible and understandable for a select group of people. As most people are left in the dark about their functioning, algorithms can be (stealthily) used to the benefit of those who control them. Without enforceable audit or other oversight mechanisms, decision-makers can obfuscate how and under what conditions they differentiate between people. Second, those with infrastructural and/or decisional power often have control over the personal information of decision-subjects. Whereas raw data may already provide decision-makers insight into decision-subjects' personal details, using inferential analytics techniques, additional knowledge (or secondary data) can be derived about decision-subjects, such as their preferences, habits, opinions or other (sensitive) attributes.[146]

Crawford, *Atlas of AI: Power, Politics, and the Planetary Costs of Artificial Intelligence* (Yale University Press, 2021), chap. 4, https://doi.org/doi:10.12987/9780300252392. See also Chapter 6, Section Algorithmic Oppression.

[144] See also: Helen Nissenbaum, 'Computing and Accountability', n.d., 9.

[145] Pasquale, *The Black Box Society*.

[146] Wachter and Mittelstadt, 'A Right to Reasonable Inferences'. E. Fosch-Villaronga et al., 'A Little Bird Told Me Your Gender: Gender Inferences in Social Media', *Information Processing & Management* 58, no. 3 (1 May 2021): 102541, https://doi.org/10.1016/j.ipm.2021.102541. Unfortunately, in certain cases, inferential analytics might be used in

Furthermore, this degree of control can be exercised over a considerable number of decision-subjects.

1.4 Conclusion

Technology and society are entangled in ways that cannot be (easily) decoupled from one another. Algorithmic decision-making systems have become an integral component of our social living environment.[147] They co-shape and co-govern people's being in the world, and alter the perception, actions, experiences, relationships and opportunities of those who interact with them.[148] The inequalities algorithms generate are an inherent part of this mediating process, and therefore inseparably linked to the societal changes these technologies might initiate.

When compared to decisions based solely on human reasoning, algorithmic techniques have given decision-makers an increased capacity to discover and/or apply both existing and/or hitherto undiscovered knowledge. Complex types of knowledge can be applied more efficiently and at a larger scale. Within the algorithmic environment, certain actors moreover hold considerable control over the means of algorithmically driven knowledge production, application, and interpretation. In combination with the amplifying properties algorithms already exhibit, these positions of centralised power add to their restructuring potential.

Though algorithms take meaning, they also shape meaning. Algorithms take meaning from their environment. They incorporate the values of their surroundings: they take and learn from data and information from the outside world, they follow the instructions of their designers, and adapt their functioning when they interact with end-users and decision-subjects. The decisions

---

ways that are outright pseudoscientific, such as algorithms that are allegedly able to detect "sexual orientation": Sam Levin, 'New AI Can Guess Whether You're Gay or Straight from a Photograph', *The Guardian*, 7 September 2017, sec. Technology, https://www.theguardian.com/technology/2017/sep/07/new-artificial-intelligence-can-tell-whether-youre-gay-or-straight-from-a-photograph; Heather Murphy, 'Why Stanford Researchers Tried to Create a "Gaydar" Machine', *The New York Times*, 9 October 2017, sec. Science, https://www.nytimes.com/2017/10/09/science/stanford-sexual-orientation-study.html; Yilun Wang and Michal Kosinski, 'Deep Neural Networks Are More Accurate than Humans at Detecting Sexual Orientation from Facial Images.', 15 February 2017, https://doi.org/10.17605/OSF.IO/ZN79K.

[147] See for instance also: Shannon Vallor, *Technology and the Virtues: A Philosophical Guide to a Future Worth Wanting* (New York: Oxford University Press, 2016), https://doi.org/10.1093/acprof:oso/9780190498511.001.0001.

[148] Verbeek, 'Materializing Morality'. Here one could also highlight the known problem of automation bias; the tendency of humans to follow overly or rely upon the results produced by automated systems.

algorithmic systems put out are therefore not neutral, but value laden. A prejudiced data set will produce prejudiced results. Yet, algorithmic systems also shape additional meaning. For instance, they do not simply replicate prejudice. Due to their scale and scope, they *reinforce* prejudice. Moreover, when decision-makers rely upon novel, hitherto undiscovered decisional grounds, they might introduce novel forms of injustice.[149] When inequality is linked to decisional grounds that appear unintuitive or incomprehensible, we might not have the requisite vocabulary to articulate why these inequalities and their social impact strike us as unjust. If we want to successfully operationalise the concept of equality for the evaluation and regulation of algorithmic decision-making systems, our notion should have the capacity to articulate and identify both types of harms.

[149] Roberge and Melançon, 'Being the King Kong of Algorithmic Culture Is a Tough Job after All', 307–8. See also: Verbeek, 'Materializing Morality'.

2 Equality as a Multidimensional Concept and the Risk of Algorithmic Egalitarian Harm

This research projects looks to revitalise the guiding role of the concept of equality for the evaluation and regulation of algorithmically guided decision-making systems. In this Chapter, I develop an understanding of equality through which we can identify and articulate the egalitarian injustice algorithms risk to introduce. In first step, and informed by political philosophy and legal theory, I examine three functions or ideals typically ascribed to the notion of equality when it is put forward as a social value guiding the structure of society. They form the basis for my understanding of the notion of equality: a *normative* concept that can be interpreted or *conceptualised* alongside three interrelated dimensions or axes; a moral, social and distributive dimension **(Section 2.1)**.[150] Though I operationalise this multidimensional concept of equality as a practicable, analytical and interpretative normative lens throughout this dissertation, in this Chapter, I use it to identify and articulate the egalitarian injustice algorithmically guided decision-making processes risk to impose onto decision-subjects.[151] To that end, the multidimensional model is positioned against the algorithmic research environment, and the practical functioning of algorithms **(Section 2.2).** Having identified the risk of egalitarian harm linked to the deployment of automated decision-making practices, I turn to the question whether in light of these risks, limitations ought to be placed on the protective scope of equality *ratione personae* in the realms of justice or the law. This question concerns the "equality of whom?" or, whom we see as the recipient or target of equality.[152]

---

[150] The idea of defining equality as a multidimensional concept is not new but fits within a certain tradition. See for instance: Axel Honneth, *The Struggle for Recognition: The Moral Grammar of Social Conflicts*, trans. Joel Anderson, Studies in Contemporary German Social Thought (Cambridge, MA, USA: MIT Press, 1996); Nancy Fraser, 'Social Justice in the Age of Identity Politics. Redistribution, Recognition, Participation.', WZB Discussion Paper, no. FS I 98-108 (1998): 29; Samuel Scheffler, 'What Is Egalitarianism?', *Philosophy & Public Affairs* 31, no. 1 (2003): 5–39, https://doi.org/10.1111/j.1088-4963.2003.00005.x; Sandra Fredman, 'Substantive Equality Revisited', *International Journal of Constitutional Law* 14, no. 3 (July 2016): 712–38, https://doi.org/10.1093/icon/mow043; Michael Walzer, *Spheres of Justice: A Defense of Pluralism and Equality* (Basic Books, Inc., 1983).

[151] As an example of such associated notions, reference can be made to domination and oppression, which are seen as inimical of equality.

[152] Iris Marion Young, 'Equality of Whom? Social Groups and Judgments of Injustice', *Journal of Political Philosophy* 9, no. 1 (1 March 2001): 1–18, https://doi.org/10.1111/1467-9760.00115.

In the final Sections, I further contextualise the notion of equality. First, I explain the relationship between equality and the related notions of bias, discrimination, and intersectionality - often seen in tandem with equality in discussions concerning algorithmic justice **(Section 2.3)**. Second, I highlight the evaluative advantage the concept of equality can offer when we assess algorithmic decision-making systems **(Section 2.4).** The Chapter concludes by explaining how the model of equality will be used in the following Parts.

2.1 Equality: A Multidimensional Concept

The notion of equality can be discussed at varying degrees of abstraction. At the highest level of abstraction, we can refer to equality to as a *concept,* whereas the *conception or conceptualisation* of equality concerns a specific interpretation given to that concept.[153] In abstract terms, a society's commitment to equality reflects its intent to equalise the enjoyment of certain fundamental interests for all. Those who favour equality of *in abstracto* can disagree in their conception or conceptualisation of equality *in concreto*.[154] For example, proponents of equality of primary goods and equality of capability both believe that people should have equal access to certain justice-relevant goods, but they nonetheless differ in terms of what they believe people should have equal amounts of.[155] In that respect equality remains, as observed by Gallie, an "essentially contested concept"; a concept, "the proper use of which inevitably involves endless disputes about [its] proper use on the part of their users."[156]

---

[153] The distinction I maintain between the notions "concept" and "conception" are inspired by the difference between "concept" and "conception" introduced by Ronald Dworkin. According to Dworkin, the distinction between the notions "concept" and "conception" concerns "a contrast between levels of abstraction." Dworkin adds: "At the first level of agreement collects around discrete ideas that are uncontroversially employed in all interpretations; at the second the controversy latent in this abstraction is identified and taken up." In my abstract concept, the dimensions I propose are not always "uncontroversial": theorists disagree as to whether equality has a socio-relational or distributive dimension. Nonetheless, they are ideal functions typically ascribed to equality. Ronald Dworkin, *Law's Empire* (Harvard University Press, 1986), 70–72 and 74; Ronald Dworkin, *Taking Rights Seriously* (Cambridge: Harvard University Press, 1977), 134–36 and 226. For an application of this distinction onto the notion of equality, see also: R. H. M. Pierik, 'The Ideal of Equality in Political Philosophy', in *The Importance of Ideals*, ed. Wibren van der Burg and Sanne Taekema (Brussels, Oxford, etc.: Peter Lang, 2004), 173–96, https://repository.ubn.ru.nl/handle/2066/67454. See also on the notion concept: W. B. Gallie, 'IX.—Essentially Contested Concepts', *Proceedings of the Aristotelian Society* 56, no. 1 (1 June 1956): 167–98, https://doi.org/10.1093/aristotelian/56.1.167; H. L. A. Hart, *The Concept of Law*, Clarendon Law Series (Oxford, England: Clarendon Press, 1961).

[154] Dworkin, *Taking Rights Seriously*; Dworkin, *Law's Empire*; Pierik, 'The Ideal of Equality in Political Philosophy'.

[155] Pierik, 'The Ideal of Equality in Political Philosophy'.

[156] Gallie, 'IX.—Essentially Contested Concepts', 169. See also. Westen, *Speaking of Equality*, 120.

In this Chapter, I advance a three-dimensional model of equality which can be further conceptualised or interpreted in function of three related, yet distinct core dimensions. The moral dimension acts as a foundational premise, which any given framework, law or policy can either accept or denounce. More specifically, a society's affirmation of moral equality entails its commitment to the core belief that every person is of equal moral worth and should be treated with equal concern and respect.[157] The socio-relational dimension relates to the level of importance a society affords in making sure that people can stand in equal relations towards one another. Finally, the distributive dimension concerns a society's commitment to equalise people's access to, or enjoyment of, certain justice-relevant goods that are needed to secure their fundamental interests.

Each dimension represents a specific function or (policy) ideal ascribed to equality as a societal value that guides the desired structuring of people's living environment. Through their interpretation of these dimensions, law, policies, and theories give normative substance to the concept of equality. This results in a specific *conceptualisation* of equality, which determines the function and importance equality holds as a governing value within a given society or framework.[158] Whereas the commitment to moral equality is binary (one accepts it or not), a society's approach towards the social and distributive dimensions can be more nuanced and gradual. In fact, a law, policy, or theory can choose to emphasise the social or distributive components of equality in varying degrees. In other words, various competing views exist

---

[157] Stefan Gosepath, 'Equality', in *The Stanford Encyclopedia of Philosophy*, ed. Edward N. Zalta, Spring 2011 (Metaphysics Research Lab, Stanford University, 2011), https://plato.stanford.edu/archives/spr2011/entries/equality/.

[158] Whereas the notion of concept is descriptive, the notion of ideal is more normative in nature. I construe equality as a three-dimensional concept. While the dimensions carry within them the potential of a given ideal, they primarily describe a set of interests typically ascribed to equality. Once these dimensions are interpreted however, they reflect an actual ideal: some type of equality we should try to achieve in practice. See: Wibren van der Burg, 'The Morality of Aspiration: A Neglected Dimension of Law and Morality', in *Rediscovering Fuller: Essays on Implicit Law and Institutional Design*, ed. Willem Witteveen J. and Wibren van der Burg (Amsterdam: Amsterdam University Press, 2009), 177–78. And Pierik, 'The Ideal of Equality in Political Philosophy', 196. For the use of the notion "ideal" to articulate equality, see Scheffler, 'What Is Egalitarianism?', 2003. For the use of the notion "dimension" as a ways to articulate equality, see: Fredman, 'Substantive Equality Revisited'. In the case of Fredman however, the notion "dimension" seems more closely aligned with that of a "normative" ideal.

concerning the way equality can be conceptualised, and the substantive goals or ideals equality aspires to achieve.[159]

The following sections are structured to reflect the fundamental and dual purpose the three dimensions have. In this first section, I will highlight the function and purpose each dimension has for structuring society. In the next section, I will assess how each dimension can inform the evaluation of algorithmically guided decision-making processes, and in this Chapter more specifically, help *identify* the egalitarian harms such systems risk to impose onto decision-subjects.

2.1.1 The Moral Dimension

In his seminal paper, "In Defense of Equality", Dworkin advanced a moral interpretation of equality, which he believed served as an egalitarian plateau; a shared premise that underlies all modern political philosophical traditions of thought regarding justice: the idea that within a society the interests of the members of that society matter, and that they all matter equally.[160] Moral equality relates to the idea that people *should be treated as equals*, which means that they should be treated with *equal concern and respect*.[161] Generally speaking, the moral dimension can be understood as concerning a deep and foundational aspect of equality, which clusters a variety of distinct but closely related ideals, including basic equality, equal concern and respect, and even

[159] A society can also decide to assign no value to equality whatsoever, or believe that other values compete with equality, and should be given priority. If equality holds no importance at all, one could say that the dimensions are empty.

[160] Will Kymlicka, ed., *Justice in Political Philosophy: Volume 1* (Cambridge: Edward Elgar Publishing Limited, 1992), xiii; Ronald Dworkin, 'Comment on Narveson: In Defense of Equality', *Social Philosophy and Policy* 1, no. 1 (1983): 24. Dworkin's "egalitarian plateau" is not shared by everyone however. See for instance: James P. Sterba, 'Recent Work on Alternative Conceptions of Justice', *American Philosophical Quarterly* 23, no. 1 (1986): 1–22. In addition, arguments for and against the inclusion of "moral equality" as a foundational principle can be found in: Uwe Steinhoff, ed., *Do All Persons Have Equal Moral Worth? On 'basic Equality' and Equal Respect and Concern*, First edition (Oxford: Oxford University Press, 2015). Scheffler and Freeman for example have pointed out the deficits of interpreting of utilitarianism as a theory of equality, although it can rephrased in "moral egalitarian" terms. Scheffler, 'What Is Egalitarianism?', 2003, 17. In reference to: Samuel Freeman, 'Utilitarianism, Deontology, and the Priority of Right', *Philosophy & Public Affairs* 23, no. 4 (1994): 313–49. However,

[161] Dworkin, *Taking Rights Seriously*; Gosepath, 'Equality'.

human dignity.[162] Moral equality, or equal concern, Dworkin argues, is the sovereign virtue of a political community, without which a government can never be considered legitimate.[163]

Policies, laws, or theories will differ in their interpretation of people's interests and the assumption that these interests matter equally.[164] They will give concrete expression to the ideal that people are one another's equals through their interpretation of the socio-relational and distributive dimension. For some, a commitment to moral equality demands the abolishment of structures of domination and oppression. For others, it requires that people be given real opportunities to pursue valuable life functionings. At the same time, the belief that people are one another's moral equals is not universally shared across societies, governments, or people.[165] Hence, from an evaluative perspective, it is best to consider the idea that people should be treated with equal concern and respect as a separate dimension of equality, the presence of which should be investigated.[166]

---

[162] Jeremy Waldron, *One Another's Equals: The Basis of Human Equality* (Cambridge, Massachusetts ; London, England: Belknap Press: Harvard University Press, 2017), 3. I shall not be concerned with the differences between these various notions. See also: George Sher, 'Why We Are Moral Equals', in *Do All Persons Have Equal Moral Worth?* (Oxford: Oxford University Press, 2014), https://doi.org/10.1093/acprof:oso/9780198719502.003.0002; Christopher Nathan, 'What Is Basic Equality?', in *Do All Persons Have Equal Moral Worth?*, ed. Uwe Steinhoff (Oxford University Press, 2014), 1–16, https://doi.org/10.1093/acprof:oso/9780198719502.003.0001. Nathan..

[163] Ronald Dworkin, *Sovereign Virtue: The Theory and Practice of Equality*, 2nd ed. (Cambridge, Massachusetts ; London, England: Harvard University Press, 2000). All the frameworks I will analyse within this dissertation, both theoretical and legal, share moral equality (or a variation thereof) as an underlying premise. For instance, the legal principle that everyone is equal before the law can be seen as an expression of the moral ideal that people should be afforded equal concern and respect in the way the law is applied to them. Hence, in the Chapters that follow, there is less need to focus specifically on equality's moral dimension. Instead, one can directly target the social and distributive dimensions.

[164] Dworkin, 'Comment on Narveson: In Defense of Equality'. If one is willing to accept that equality has a deep moral dimension that should function as an underlying ideal - or a variation thereof – and foundational element of justice, every person would have to be considered as a subject and beneficiary of justice, and moreover, a subject and beneficiary of justice on equal terms. In that sense, moral equality is a prescriptive ideal. Waldron, *One Another's Equals: The Basis of Human Equality*, 47.

[165] Acts of racism can be motivated by a wrongful sense of superiority and actively undermine the equal moral worth of people because of their differences in culture, ethnicity, or background.

[166] Furthermore, to assign certain groups or people within society greater moral worth, does not exclude equality in a social or distributive sense. For example, in a meritorious framework moral worth may first need to be earned through good actions and behaviour, before a person can make a moral claim towards treatment as an equal. See Nathan, 'What Is Basic Equality?', 14. For more discussions concerning the notion, see for instance the various contributions in: Steinhoff, *Do All Persons Have Equal Moral Worth?*

PATROLLING AND INCLUSION - The idea that people are one another's moral equals (in some way or another) can be relevant for evaluating justice and differential treatment.[167] Waldron for instance has argued that a foundational and moral understanding of equality can "patrol and discipline" disputes concerning the existence of social, relational and material (in)equalities within society.[168] Under this view, the presence of social and distributive inequalities need not be problematic *per se*, but a society that asserts the moral equality of its citizens would need to put forward good reasons as to why the existence of social and economic inequalities are in line with the commitment that everyone's interests matter and matter equally. Moral equality is therefore also inclusionary. If people are one another's equals, it seems reasonable to assume that no single person should *a priori* be excluded from the specific functioning mechanisms of a given policy or framework.[169] Again, the existence of socio-relational and distributive inequality would not be precluded, but the presence of these inequalities should be the result of a system where anyone could *ab initio* participate on equal terms. In this regard, moral equality can be interpreted as being a procedural safeguard which protects people against the arbitrary application of a given policy or ruleset.[170] Moral equality thus becomes a type of review mechanism to evaluate a framework's consistency and coherence in relation to its alleged commitment to equality: when a framework or policy claims that people's interests were equally considered, how serious were they in their actual implementation of this ambition?[171]

---

[167] Kymlicka points towards an additional advantage of a shared "moral egalitarian" premise among theories: it would facilitate the comparison of those frameworks that share it as an underlying premise. Kymlicka, *Justice in Political Philosophy: Volume 1*, xiii. For another example on how equality's moral dimension can be used as a bridge, see also: Pierik, 'The Ideal of Equality in Political Philosophy'.

[168] Waldron, *One Another's Equals: The Basis of Human Equality*, 14.

[169] For many frameworks this commitment appears (theoretically) confirmed: every person has a vote in the utilitarian calculus, every person has a freedom to contract in a neoliberal free market, every person has access to a basic threshold of central capabilities following a capability approach, etc. Certain theories have been criticized however from not allowing such participation. See for example Nussbaum's critique of Rawls' Justice as Fairness vis-à-vis disabled people by targeting "rational actors."

[170] As a procedural ideal, equality governs the *equal or consistent application of a given procedure or rule.* Once a specific ruleset has been adopted, procedural equality guarantees that no one is arbitrarily excluded from being subject to the application of the rules as they were agreed upon.

[171] This patrolling function can also be used to guard against counter-intuitive interpretations of moral equality. See for example Dworkin's critique of utilitarianism's egalitarian foundations, and Pettit's analysis thereof. For instance, a utilitarian calculus where every person's vote counts as one, and not more than one, in principle, satisfies the conditions for moral equality: every person is treated with equal concern and respect in the sense that each person has an equal say and sway in the outcome of the voting system. The adoption of a sexist hiring policy, even in the case it would lead to grave socio-relational and socio-economic inequalities for women, would be entirely justified if women were given an equal vote, but found themselves simply outnumbered. The alleged commitment moral equality leads to a minority's subjugation to inferiorising treatment (socio-relational harm) and loss of valuable socio-economic opportunities

2.1.2 The Socio-Relational Dimension

A just society recognises people's equal moral status not only in theory, or vis-a-vis the distribution of certain (non)material goods, but also more profoundly in their day-to-day interactions and relations. A just society is a community of free and equal persons where every person is respected as a valued member of that community.[172] The fundamental interest is one of relational equality. The socio-relational dimension of equality reflects the lengths at which a society, law or policy is willing to go in order to ensure people's recognition as equals in the various relationships they hold.[173]

Social or relational equality looks to govern people's relationship in a two-fold way. On the one hand, social egalitarians will formulate principles and propose practices that structure society,

---

(distributive harm). The initial equal concern and respect given to people, by affording everyone a vote in the utilitarian calculus, is defeated by allowing people to vote for a system where people are not treated with equal concern and respect. The pertinent question one can raise is whether in the case a system's alleged commitment to moral equality ends up being self-defeating , should that system be considered as a preferred governing structure? Pettit, *Judging Justice*, 131–32.

[172] This follows Schuppert's observation, who notes that "The basic normative assumption from which champions of social equality set out is the idea of a just society as a community of free and equal persons." Fabian Schuppert, 'Non-Domination, Non-Alienation and Social Equality: Towards a Republican Understanding of Equality', *Critical Review of International Social and Political Philosophy* 18, no. 4 (4 July 2015): 444, https://doi.org/10.1080/13698230.2015.1033863. Other relevant contributions exploring the meaning of social equality: Carina Fourie, Fabian Schuppert, and Ivo Walliman-Helmer, eds., *Social Equality: On What It Means to Be Equals* (New York: Oxford University Press, 2015); Schuppert, 'Non-Domination, Non-Alienation and Social Equality'; Scheffler, 'What Is Egalitarianism?', 2003; Elizabeth S. Anderson, 'What Is the Point of Equality?', *Ethics* 109, no. 2 (1999): 287–337, https://doi.org/10.1086/233897.

[173] The question whether moral and social equality are distinguishable from one another has been a point of contention. Fourie and others note that a social egalitarian may therefore investigate whether social equality is "an expression of equal moral worth or treating people with respect and concern", or whether "it is a substantive ideal in itself that needs to be fleshed out , and this fleshing out will help to determine which distributive principle are compatible with equality." Carina Fourie, Fabian Schuppert, and Ivo Walliman-Helmer, 'The Nature and Distinctiveness of Social Equality: An Introduction', in *Social Equality: On What It Means to Be Equals*, ed. Carina Fourie, Fabian Schuppert, and Ivo Walliman-Helmer (New York: Oxford University Press, 2015), 5–6. I believe the social dimension is worth highlighting as a separate ideal, as it may paint a more nuanced and intricate vision concerning the various functions equality may hold. See *among others:* Anderson, 'What Is the Point of Equality?', 288–89; Marie Garrau and Cécile Laborde, 'Relational Equality, Non-Domination, and Vulnerability', in *Social Equality: On What It Means to Be Equals*, ed. Carina Fourie, Fabian Schuppert, and Ivo Walliman-Helmer (New York: Oxford University Press, 2015), 47. Discussions concerning the relational nature of equality may also pertain the need to have a "comparator." However, I shall not be concerned with this question here. For the latter, see also: Gosepath, 'Equality'.

and the actions and behaviours of citizens, in such a way that people can stand in relations of equality.[174] For instance, proponents of relational equality may argue that a political decision-making process must be inclusive and allow every person to raise their concerns and interests.[175] On the other hand, social egalitarians will look to identify and mitigate a wide variety of harms, and associated practices, of which they believe they are inimical to people's standing as free, equal and valued members of society.[176] These harms can take effect in the relationships people and groups of people hold to each other, as well as to themselves.[177]

A society can be characterised by class struggles, status-hierarchies, a caste system, or other power asymmetries which afford one person or group a greater degree of control over others. In such an environment, a person or group may find themselves in a position where they are rendered more vulnerable to various forms of injustice imposed by others: they can be dominated, oppressed, stigmatised, etc. [178] At the same time, the presence of these unequal social relations can have an impact on the ways people behave, act and perceive themselves: they can feel inferior, lose their sense of self-respect, act in ways that are more subservient, be anxious due to other people's capacity to interfere in their lives, etc.[179] Hierarchies in power need not be forbidden, but a society that is concerned with social equality will make sure that those powers are exercised responsibly.[180] Laws and policies will differ in the harms they consider most antagonistic towards people's relationships, and they will also differ in the conditions they put in place in order to preserve social equality.[181]

---

[174] Anderson, 'What Is the Point of Equality?', 313; Samuel Scheffler, 'The Practice of Equality', in *Social Equality: On What It Means to Be Equals*, ed. Carina Fourie, Fabian Schuppert, and Ivo Wallimann-Helmer (New York: Oxford University Press, 2015), 21–22.

[175] Iris Marion Young, *Inclusion and Democracy* (Oxford University Press, 2002), https://doi.org/10.1093/0198297556.001.0001.

[176] Schuppert, 'Non-Domination, Non-Alienation and Social Equality'; in reference to Martin O'Neill, 'What Should Egalitarians Believe?', *Philosophy & Public Affairs* 36, no. 2 (March 2008): 119–56, https://doi.org/10.1111/j.1088-4963.2008.00130.x.

[177] Schuppert, 'Non-Domination, Non-Alienation and Social Equality', 445.

[178] Anderson, 'What Is the Point of Equality?'; Young, *Justice and the Politics of Difference*. See also Scheffler, 'What Is Egalitarianism?', 22. In interpreting Anderson. Anderson, 'What Is the Point of Equality?'

[179] Schuppert, 'Non-Domination, Non-Alienation and Social Equality', 445.

[180] Conversely, where a person or group are able to exercise their power over others in ways that are abusive, social equality is not pursued or insufficiently protected. See also: Fourie, Schuppert, and Walliman-Helmer, *Social Equality: On What It Means to Be Equals*, 3.

[181] They may also differ in who should be the target of a given measure: individuals, social groups, etc.

2.1.3 The Distributive Dimension

Characteristic of the distributive paradigm is that it formulates questions of justice as questions of distribution. Distributive justice considers whether a society's political, social and economic institutions distribute the opportunities, burdens, and benefits people carry in a way that best accommodates their interests.[182] According to Young, those who follow a distributive paradigm, believe any social value within society can be explained in terms of the possession of certain material (e.g., wealth and income) or non-material goods (e.g., opportunities and social positions).[183] Distributive justice is concerned with those justice-relevant goods people are entitled to receive.[184] The goods and positions people hold are seen as representative of people's overall interests and any interference in their distribution, is therefore an interference, negative or positive, in people's overall interests.

In this context, equality becomes a (purely) distributive value or ideal when justice mainly requires for people to have equal access to, or equal amounts of, something.[185] Whereas equality as a social ideal primarily looks for ways to govern and equalise the relations people stand in, distributive egalitarians tend to define the demands of justice through the identification of those

---

[182]Julian Lamont and Christi Favor, 'Distributive Justice', in *The Stanford Encyclopedia of Philosophy*, ed. Edward N. Zalta, Winter 2017 (Metaphysics Research Lab, Stanford University, 2017), https://plato.stanford.edu/archives/win2017/entries/justice-distributive/. Pettit, *Judging Justice*. David Miller, 'Justice', in *The Stanford Encyclopedia of Philosophy*, ed. Edward N. Zalta, Fall 2017 (Metaphysics Research Lab, Stanford University, 2017), https://plato.stanford.edu/archives/fall2017/entries/justice/. As the distribution of resources necessitates a certain institutional structure, Barry considers distributive justice to be institutional in nature. In this sense, justice becomes an attribute of institutions. Under this view, justice stands for people's value judgments concerning an institution's chosen distribution of benefits and burdens across people. Brian Barry, *Theories of Justice: A Treatise on Social Justice, Volume I* (Berkeley and Los Angeles, California: University of California Press, 1989), 355.

[183] Young herself has critiqued this purely distributive understanding of justice. By focusing on end-state patterns, she argues that the application of the distributive paradigm onto non-material and social goods obscures how the social and economic position of people comes about as a result of social processes. See also: Young, *Justice and the Politics of Difference*, 24–25.

[184] Schemmel, 'Distributive and Relational Equality', 126–27; Thomas W. Pogge, 'Three Problems with Contractarian-Consequentialist Ways of Assessing Social Institutions*: THOMAS W. POGGE', *Social Philosophy and Policy* 12, no. 2 (1995): 241–66, https://doi.org/10.1017/S026505250000474X.

[185] Scheffler, 'The Practice of Equality', 21. In reference to Cohen: G. A. Cohen, 'On the Currency of Egalitarian Justice', *Ethics* 99, no. 4 (1989): 906. See also: Amartya Sen, 'Equality of What?' (The Tanner Lecture on Human Values, Stanford University, 1979), 26; Schemmel, 'Distributive and Relational Equality'; Kristin Voigt, 'Distributive Equality, Relational Equality and Preferences about Higher Education', *Theory and Research in Education* 15, no. 2 (1 July 2017): 109–28, https://doi.org/10.1177/1477878517708116.

things (justice-relevant goods) that are so fundamental to person's life that each person should have an equal entitlement thereto, and advance principles in an effort to secure or maintain this entitlement.[186]This debate is often formulated as looking for the *currency* of egalitarian justice or, in the words of Amartya Sen, the "equality of what?"[187] This currency is a criteria or metric by which the situation of persons and the inequalities they incur, can be evaluated and compared.[188] Over the years, many so-called currencies of egalitarian justice have been proposed and scrutinised: equality of opportunity of welfare[189]; equality of primary goods[190]; equality of resources[191]; equality of capability[192]; and so forth.[193] In an effort to assess whether inequalities can be justified, distributive egalitarians will measure the advantage and disadvantage in relation to the identified currency of justice.[194] That being the case, the individual shares people hold of those things that contribute to the identified fundamental interest are of major importance.[195] Put differently, as a distributive notion, equality pertains to differences in people's outcomes, such as their opportunities and welfare, and when we look for justice, we evaluate whether those outcome inequalities are justified or not.[196]

---

[186] And if necessary, formulate rules that can allow for a deviation of those standard principles. Schemmel, 'Distributive and Relational Equality', 126.

[187] The notion currency or metric was made famous by G.A. Cohen in Cohen, 'On the Currency of Egalitarian Justice'. Cohen's influential article was a response to Amartya Sen, who phrased the inquiry into the metric of egalitarian justice as the question concerning the "equality of what?": Amartya Sen, *The Idea of Justice* (Belknap Press: Harvard University Press, 2009). See also: Ingrid Robeyns, 'The Capability Approach', in *The Stanford Encyclopedia of Philosophy*, ed. Edward N. Zalta, Winter 2016 (Metaphysics Research Lab, Stanford University, 2016), https://plato.stanford.edu/archives/win2016/entries/capability-approach/.

[188] Sen, 'Equality of What?'; Cohen, 'On the Currency of Egalitarian Justice'; Richard J. Arneson, 'Welfare Should Be the Currency of Justice', *Canadian Journal of Philosophy* 30, no. 4 (December 2000): 497–524, https://doi.org/10.1080/00455091.2000.10717541; Gosepath, 'Equality'; Schemmel, 'Distributive and Relational Equality'. Scheffler, 'The Practice of Equality', 21; Cohen, 'On the Currency of Egalitarian Justice'.

[189] Richard J. Arneson, 'Equality and Equal Opportunity for Welfare', *Philosophical Studies: An International Journal for Philosophy in the Analytic Tradition* 56, no. 1 (1989): 77–93.

[190] Rawls, *A Theory of Justice*, 1971.

[191] Ronald Dworkin, 'What Is Equality? Part 2: Equality of Resources', *Philosophy & Public Affairs* 10, no. 4 (1981): 283–345.

[192] Sen, 'Equality of What?'; Sen, *The Idea of Justice*; Martha C. Nussbaum, *Creating Capabilities: The Human Development Approach* (Cambridge, Massachusetts ; London, England: The Belknap Press of Harvard University Press, 2011).

[193] Policies may agree on the currency in question, but not on the means that need to be put in place to achieve said metric.

[194] Schemmel, 'Distributive and Relational Equality', 127.

[195] Schemmel, 127.

[196] Voigt, 'Distributive Equality, Relational Equality and Preferences about Higher Education', 112.

Distributive egalitarians do not preclude the existence of socio-economic inequality, nor do they advance a type of strict or simple equality. Under a strict egalitarian regime, every single person would be given the same share of something, and a just society would strive towards an end-state where every person is given the same material level of goods of services.[197] Yet, this type of equal distribution is hardly ever argued for, and would moreover act counterproductive towards the achievement of most mainstream egalitarian aims.[198] For instance, if we want to guarantee that every person has an equal opportunity to enter the job market, it makes sense to allocate more (institutional) resources to persons that typically face hardship in looking for a job. The realisation of such policy goals is not excluded when following a distributive paradigm. In similar fashion, under certain frameworks, equality will be achieved as soon as people have a "sufficient" amount of something.[199]

2.1.4 Dimensional Interdependence

The above-mentioned dimensions were presented as distinct. Understood as a purely relational ideal, equality remains keenly centred on the relationships people hold, and in the way people ought to treat one another as part of these relationships.[200] Unlike conceptualisations of equality that are distributive in nature, the main concern lies not in what people should have access to in terms of (non-)material goods. Instead, the social egalitarian will look for a distributive ruleset (among others) that best fits the relational ideal: what do people need in order stand in relations

---

[197] Gosepath, 'Equality'. Strict equality thus acts as a type of procedural equality: every person must be given the same.

[198] For instance, if money were to be the central distributable good, people with a physical disability may benefit from having a monetary compensation similar to a doctor's wage, but their participation in society would remain dependent upon non-monetary initiatives, such as measures that enhance accessibility, such as the presence of wheelchair ramps in the public space. Similarly, in order to ensure equality of opportunity within society, non-discrimination laws allow legislators to introduce affirmative action measures, which, in order to ensure equality on a societal level, facilitate a (short-term) unequal distribution of certain goods within society. The argument in favour of affirmative action is that in order to correct historical disadvantages in relation to access to specific goods or positions, and in order to ensure equal treatment in the future, the disadvantaged group should receive preferential treatment in the now.

[199] Moreover, even in the case it is argued that sufficiency and priority are entirely unrelated to equality as distributive mechanisms, they may nonetheless be egalitarian from a moral or social perspective. For instance, in aforementioned prioritarian view, those with more needs might be given priority in their access to certain societal benefits, but this could be motivated by the desire to ensure them the capacity to attain an equal relational status to those who are currently better off. On competing views, see also: Gosepath, 'Equality'.

[200] Schemmel, 'Distributive and Relational Equality'; John Baker, 'Conceptions and Dimensions of Social Equality', in *Social Equality: On What It Means to Be Equals*, ed. Carina Fourie, Fabian Schuppert, and Ivo Walliman-Helmer (New York: Oxford University Press, 2015), 65–86; Garrau and Laborde, 'Relational Equality, Non-Domination, and Vulnerability'.

of equality, and what protection should people be given so they are ensured treatment as equals?[201] While a social egalitarian will acknowledge the importance of both people's relationships and institutionally governed affordances, the distributional ideal of equality is subsumed and given shape by the social ideal. Under a purely distributional conceptualisation of equality, a just society guarantees a correct distributive state of affairs concerning the "justice-relevant goods" that secure people's fundamental interests.[202] While distributive equality might be concerned with the transformation of institutional structures or people's relationships, these transformations are seen as instrumental towards achieving the distributional ideal.

In reality however, it is likely that the various dimensions share - or should share - a relationship of interaction and reciprocity.[203] In a society where equality is valued as both a social and socio-economic, structuring ideal, one would expect to find a strong socio-relational conceptualisation of equality at its core, supported by a robust distributive framework.[204] In certain cases, a theory, law or framework will indeed be highly interdependent, whereby each respective dimension fulfils a core and integral function towards the pursuit of an overall goal or aim.[205] For instance, in

---

[201] Scheffler, 'The Practice of Equality', 22. On the relationship between social and distributive justice: Young, *Justice and the Politics of Difference*; Fraser, 'Social Justice in the Age of Identity Politics. Redistribution, Recognition, Participation.'; Anderson, 'What Is the Point of Equality?'; Sandra Fredman, 'Redistribution and Recognition : Reconciling Inequalities', *South African Journal on Human Rights* 23, no. 2 (January 2007): 214–34, https://doi.org/10.1080/19962126.2007.11864923. Garrau and Laborde, 'Relational Equality, Non-Domination, and Vulnerability', 48. Moreover, as Anderson has pointed out, there is an interplay between people's social context and their distributive needs: "What counts as "enough" varies with cultural norms, the natural environment, and individual circumstance." Anderson, 'What Is the Point of Equality?', 320.

[202] Schemmel, 'Distributive and Relational Equality', 127. In reference to Pogge: Thomas Pogge, 'Relational Conceptions of Justice: Responsibilities for Health Outcomes', in *Public Health, Ethics, and Equity*, ed. Sudhir Anand, Fabienne Peter, and Amartya Sen (Oxford University Press, 2004), 147.

[203] As Schemmel observes: "An integrated relational view that managed to unite these two dimensions under an overarching framework of principles of social justice would be a theory with greater explanatory power." Schemmel, 'Distributive and Relational Equality', 129.

[204] For an example on the reconciliation between the socio-relational and distributional dimensions of equality: Nancy Fraser and Axel Honneth, *Redistribution Or Recognition?: A Political-Philosophical Exchange* (Verso, 2003); Fraser, 'Social Justice in the Age of Identity Politics. Redistribution, Recognition, Participation.'

[205] The dimensional interpretations of equality may also be influenced by views on personhood. What does it mean for a person to lead a (dignified) life in terms of how they should be treated from a moral and social perspective? What goods should a person have access to aid them therein? Do people have universal interests, behavioural characteristics, traits, and abilities that must be protected? The answers to these questions bring together a specific vision on personhood and can further explain the interconnectedness between a given framework's moral, social and distributive dimensions. For instance, libertarians see people as rational self-owners that possess strong natural rights. Because the decisional power people hold is seen as sacred, collective interests and social inequalities are less problematic. In fact,

order to prevent social exclusion, policy makers might promote equal access to educational opportunities.

In practice this co-dependence between social and distributive equality does not always exist. It therefore is still valuable to make a distinction between the dimensions from an evaluative perspective. Regulators for instance can have different opinions as to the types of equality a given law or policy should pursue, and how they should be pursued.[206] Furthermore, the role of equality can be linked to, or identified through, other values, such as dignity, respect, reciprocity or autonomy.[207] Racism for example can be seen as both a socio-relational and dignitarian harm. To deny a person or group of people the respect they deserve because of their skin colour, ethnicity or nationality is to treat them as inferior and of lesser social standing. Likewise, if one group is denied access to an important public resource, such as education, their capacity towards self-development and self-determination is limited.[208]

---

welfare and taxation policies are seen as an impermissible interference with a person's right to pursue their own life's goals. Capability theorists are more sensitive towards the societal conditions under which people grow up. People may be held responsible for their choices, but only as far as choice is freely developed and exercised. Both frameworks perceive people differently in terms of the responsibilities they bear, and this outlook on responsibility affects what inequalities they find justified. A theory's construction of personhood may therefore also affect the evaluation of algorithmic decision-making systems, including the inequalities such systems generate. For example, one could argue that, under a capability view, covert or manipulative profiling techniques actively interfere with people's agency and control. As a consequence, the inequalities people incur when they enter into such a system could not be justified.

[206] Policy makers may believe it is important to recognise the precarious socio-relational position of marginalised and historically disadvantaged groups but consider that their recognition is best secured through a series of socio-economic policies promoting equality of opportunity in education and employment, rather than sensibilisation campaigns that look to convince the population at large that every person has equal moral worth. At the same time, equality may not be given a prominent role whatsoever. A society may prioritise individual liberty above everything else. Here, equality is not a social or distributional ideal that is protected through a given policy choice. Instead, if equality (or non-discrimination) is to exist it all, it is the by-product of people's free choices.

[207] See also: Scheffler, 'The Practice of Equality', 41. Scheffler believes that the "distinctive egalitarian aspects" of the socio-relational view "cannot be identified, nor their appeal appreciated without reference to other values."

[208] For a more in-depth elaboration of the notions of self-development and self-determination and how they are pressured within the algorithmic environment, reference can be made to Chapter 6, Section 6.1. "Algorithmic Oppression and Domination." Drawing from Young, the right to self-development is understood as a person's capacity to learn, develop, exercise, and express their skills, capacities, and experiences in socially meaningful and recognised ways. The right to self-determination is defined as a person or group's capacity to exercise control, among others through participation, over their own options and choices and the conditions of their actions.

Much more can be said about the theoretical and conceptual distinction between the identified dimensions and how they represent differing approaches towards equality.[209] More interesting however, and my core concern, is to assess how each dimension and ideal equality represents can be linked to the various operations and purposes for which algorithms are used in order to identify the potential injustices or egalitarian harms algorithms risk to impose.

2.2 Algorithmic Inequality: The Risk of Egalitarian Harm

2.2.1 Algorithmic Egalitarian Harms

In this Section, I will illustrate how algorithms, through their functioning, risk to impose egalitarian harms onto decision-subjects. Taking the distinction between algorithms as meaning-takers and meaning-givers, this dissertation focuses on two broad types of injustice.[210] On the one hand, algorithms tend to perpetuate and reinforce structural (and often systemic) inequalities faced by historically disadvantaged groups. On the other hand, algorithms by their nature risk to produce egalitarian harms more broadly and indiscriminately.

This Section will first explain how the inequalities that are often experienced by historically disadvantaged groups in the digital environment can be reformulated as moral, socio-relational, and distributional egalitarian harms. This section then illustrates how the dimensions identified within the previous section can be successfully used to explain the particular egalitarian harms (historically) disadvantaged groups are said to face within the algorithmic environment. In the paragraphs that follow however, those dimensions will be relied upon to define egalitarian harms more broadly, namely as harms algorithmic decision-making systems risk to impose onto everyone they (in)directly govern regardless of their current social position in life. The three dimensions highlight specific functions or (policy) ideals that have been ascribed to the concept of equality for the structuring of society. Algorithmically guided decision-making systems can interfere with these ideals: they can restructure people's relationships and they can affect the allocation of justice-relevant goods. The dimensions can therefore be mapped onto the workings

[209] For more in depth discussions, reference can be made to: Schemmel, 'Distributive and Relational Equality'; Scheffler, 'What Is Egalitarianism?', 2003; Scheffler, 'The Practice of Equality'; Pogge, 'Relational Conceptions of Justice', 2004; Fraser and Honneth, *Redistribution Or Recognition?*; Anderson, 'What Is the Point of Equality?'; Young, *Justice and the Politics of Difference*. As well as the various contributions in: Fourie, Schuppert, and Walliman-Helmer, *Social Equality: On What It Means to Be Equals*.

[210] See Chapter 1

of algorithms, and the specific questions that each dimension raises, can inform the way we approach algorithmic decision-making systems.[211]

Historical Disadvantage

Algorithmically guided decision-making systems replicate and reinforce the harms and injustice suffered by historically disadvantaged groups; marginalised groups; minorities; people who find themselves in a more vulnerable positions; or people whose struggles have been made invisible. In every society, and regardless of who the persons belonging to the aforementioned disadvantaged groups are, algorithms have imposed or reinforced systemic structures of oppression to the detriment to those not in a position of power.[212] Unfortunately, such examples are myriad.[213] Eubanks' research illustrated how the automation of welfare systems in the United States disproportionately punished the poor.[214] In the case of the Dutch SyRI (System Risk Indicator) programme, Philip Alston, the United Nations Secretary on Extreme Poverty and Human Rights criticized its predominant use in areas characterised by a higher concentration of poverty and vulnerability.[215] In other research, Buolomwini and Gebru showed how black faces, and more specifically, the faces of black women were more often falsely classified by facial recognition algorithms.[216] Automated gender recognition systems misgender trans- and non-binary persons.[217]

---

[211] In the next Chapters, I will address more concretely how such harms are effectively addressed by investigating what inequalities are deemed justified or not. First, by looking into the legal conceptualisation of equality found within European Equality and Non-Discrimination Laws. In the final Chapters, I will return to these harms by analysing how concrete conceptualisations of equality found within theories of justice could make these harms more concrete.

[212] See for example the work of D'Ignazio and Klein for an overview of how existing power infrastructures creep into the digital environment. D'Ignazio and Klein, *Data Feminism*.

[213] See also: Safiya Umoja Noble, *Algorithms of Oppression: How Search Engines Reinforce Racism*, *Algorithms of Oppression* (New York University Press, 2018), https://doi.org/10.18574/9781479833641; Cathy O'Neil, *Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy* (New York, NY, USA: Crown Publishing Group, 2016).

[214] Eubanks, *Automating Inequality: How High-Tech Tools Profile, Police and Punish the Poor*.

[215] https://www.ohchr.org/Documents/Issues/Poverty/Amicusfinalversionsigned.pdf

[216] Joy Buolamwini and Timnit Gebru, 'Gender Shades: Intersectional Accuracy Disparities in Commercial Gender Classification', in *Conference on Fairness, Accountability and Transparency* (Conference on Fairness, Accountability and Transparency, PMLR, 2018), 77–91, http://proceedings.mlr.press/v81/buolamwini18a.html.

[217] 'How AI Systems Undermine LGBTQ Identity', *Access Now* (blog), 6 April 2021, https://www.accessnow.org/how-ai-systems-undermine-lgbtq-identity/. See also: Os Keyes, 'The Misgendering Machines: Trans/HCI Implications of Automatic Gender Recognition', *Proceedings of the ACM on Human-Computer Interaction* 2, no. CSCW (November 2018): 1–22, https://doi.org/10.1145/3274357; Morgan Klaus Scheuerman, Jacob M. Paul, and Jed R. Brubaker, 'How

The reasons as to why algorithms reify or perpetuate injustice have been well-documented.[218] The inequalities experienced by historically disadvantaged groups can first be reflected within the data sets used to train an algorithm, resulting in biased decision-making procedures.[219] Second, historic forms of injustice may also be introduced (unknowingly) through choices in the design.[220] Stereotypical world views may inform the design choices system developers make. Certain of these choices may also reflect normalised modes of thinking that can nonetheless be questioned on their biological, scientific, and moral validity.[221] For instance, gender recognition systems are structured around a gender-binary and straightforward views on sex and gender that are outdated.[222] Finally, social injustice may also come about in real-time due to human interaction upon deployment. Here social inequality is a product of the intended or unintended, visible, or invisible, actions of many people.[223]

Computers See Gender: An Evaluation of Gender Classification in Commercial Facial Analysis Services', *Proceedings of the ACM on Human-Computer Interaction* 3, no. CSCW (7 November 2019): 1–33, https://doi.org/10.1145/3359246.

[218] The arrival of these injustices can often be explained through Nissenbaum and Friedman's typology of computer bias. See below Section 2.3 Bias and Inequality. Batya Friedman and Helen Nissenbaum, 'Bias in Computer Systems', *ACM Transactions on Information Systems* 14, no. 3 (1 July 1996): 330–47, https://doi.org/10.1145/230538.230561. See also: Ninareh Mehrabi et al., 'A Survey on Bias and Fairness in Machine Learning', *ArXiv:1908.09635 [Cs]*, 17 September 2019, http://arxiv.org/abs/1908.09635.

[219] For instance, women and people of colour are often less likely to show up in image search results of people looking for CEOs because they are (still) underrepresented in higher positions. Hence, there is less pictorial reference material. Even if all the images of CEOs in the world were collected, the structural underrepresentation of marginalised groups in higher positions would still be vividly visible. See: Engin Bozdag, 'Bias in Algorithmic Filtering and Personalization', *Ethics and Information Technology* 15, no. 3 (2013): 209–27, https://doi.org/10.1007/s10676-013-9321-6; Mehrabi et al., 'A Survey on Bias and Fairness in Machine Learning'; Kate Crawford, 'The Hidden Biases in Big Data', *Harvard Business Review*, 1 April 2013, https://hbr.org/2013/04/the-hidden-biases-in-big-data; Friedman and Nissenbaum, 'Bias in Computer Systems'.

[220] If more facial images are collected of white men, than of women and people of colour, facial recognition software will fail to recognize the latter categories more accurately. Buolamwini and Gebru, 'Gender Shades'.

[221] See for instance: Denton et al., 'Bringing the People Back In'; Miceli et al., 'Documenting Computer Vision Datasets'; Kate Crawford and Trevor Paglen, 'Excavating AI: The Politics of Images in Machine Learning Training Sets', Excavating AI, 19 September 2019, https://www.excavating.ai; Crawford, *Atlas of AI: Power, Politics, and the Planetary Costs of Artificial Intelligence*.

[222] Thekla Morgenroth and Michelle K. Ryan, 'The Effects of Gender Trouble: An Integrative Theoretical Framework of the Perpetuation and Disruption of the Gender/Sex Binary', *Perspectives on Psychological Science* 16, no. 6 (1 November 2021): 1113–42, https://doi.org/10.1177/1745691620902442.

[223] An extreme example concerns the Microsoft developed Twitter bot Tay. The purpose of Tay was to learn from user interaction, but 24 hours after its introduction, the robot had become a bigoted fascist. Sure, Tay may have learned from historic data too, and certain design choices might have led to Tay's transformation from good to evil, but in this particular case, social injustice was live fed through real-time human interaction. Elle Hunt, 'Tay, Microsoft's AI Chatbot,

In general, the claims made by historically disadvantaged groups towards greater algorithmic justice correspond to traditional claims for social justice. Their call towards equality is one of recognition and distribution.[224] While these concerns may take on new forms, they often pertain to old, persistent problems: failure to identify the members of certain social groups as equals (moral), both in the relationships they hold (social), and in the opportunities, burdens and benefits they are afforded (distributive).[225] By replicating the injustice social groups face in the analogue environment, the digital environment further solidifies the systemic and structural nature of the injustice social groups have experienced. The digital environment becomes yet another social layer wherein these groups face adverse treatment. The fact that algorithmic processes, which have properties that further amplify the differences between people, can perpetuate these inequalities under a veil of opacity and complexity, should be an additional reason to heighten our caution.

Gets a Crash Course in Racism from Twitter', *The Guardian*, 24 March 2016, sec. Technology, https://www.theguardian.com/technology/2016/mar/24/tay-microsofts-ai-chatbot-gets-a-crash-course-in-racism-from-twitter.

[224] Fraser, 'Social Justice in the Age of Identity Politics. Redistribution, Recognition, Participation.'; Anderson, 'What Is the Point of Equality?'; Honneth, *The Struggle for Recognition*; Axel Honneth, 'Recognition and Justice: Outline of a Plural Theory of Justice', *Acta Sociologica* 47, no. 4 (December 2004): 351–64, https://doi.org/10.1177/0001699304048668; Fredman, 'Redistribution and Recognition'.

[225] Like misgendering machines, administrative forms, such as medical information or identity cards, were often created along a gender binary, imposing representational harms onto gender non-conforming persons. Even before the arrival of racially biased policing systems, people of foreign origin were often the subject of more stop-and-checks on the streets. People living in poverty are subject to heightened digital surveillance, but they already faced heightened scrutiny in their consultations with caretakers. Not every instance of differential treatment towards a social interest group affects a historically disadvantaged group. The promotion of right-wing over left-wing ideology in the online environment or the "shadow banning" of climate activism could be unjustified, but these measures do not (directly) affect historically disadvantaged groups. Huszár et al., 'Algorithmic Amplification of Politics on Twitter'. Josephine Lulamae, 'Automated Society: Are Green Influencers Being "Shadowbanned" on Instagram?', accessed 11 February 2022,
https://r.algorithmwatch.org/nl2/x3ysn/158g.html?m=AL8AAL79d5EAAAAPjTIAAAZJGn8AAAAA6uQAAB1gABB0KQBiBfudOgjIB3G0TWetucmFgzeQTAAQJ0I&b=bfbfe00e&e=3d32336c&x=0A1rc6D5oqSxnwruhXsAaIRdCFsr9ZW1dtSNvoBxFTE. On the notion "data justice", see also: Lina Dencik and Javier Sanchez-Monedero, 'Data Justice', *Internet Policy Review* 11, no. 1 (14 January 2022), https://policyreview.info/articles/analysis/data-justice.

Algorithmic Moral Harms

A commitment to moral equality legitimises an inquiry as to whether algorithmic decision-making systems treat people with equal concern and respect. The patrolling and inclusionary function of equality's moral dimension can be used to scrutinise algorithmic practices in order to investigate whether they effectively treat people with equal concern and respect. Whenever an algorithm operates in a way that (appears to) deviate from the default premise that people's interests matter and matter equally, it gives us a motivation to investigate their functioning, development, and deployment. For example, if an algorithm differentiates between people, and in doing so (systematically) advantages one group to the detriment of other groups, we should ask ourselves the question: how exactly does this differential treatment benefit the interests of those involved, and if not, why is this okay?

In a similar vein, our commitment to moral equality can be a critical argument against biased decision-making systems.[226] Indeed, given a system's bias inevitably entails that one group or outcome will be (dis)favoured by the system over others, can it still treat every person with equal concern and respect? While decision-makers should be given an opportunity to justify their actions, moral equality can be used to shift the burden of proof. The onus would be on the algorithmic decision-maker to show that a given practice *did* take people's interests into account; or, in the case they failed to do so, or where other goals were pursued, why they were justified in not doing so.

In practice however, the question whether an algorithmic decision-making system treats people with equal concern and respect will often need to be resolved with the support of a social or distributive conceptualisation of equality. Those interpretations give more specific substance to the moral ideal of equality, and in most cases a violation of social or distributive equality, which typically signals the violation of people's moral equality, will be more easily discovered.

Algorithmic Socio-Relational Harms

Questions of social equality are intricately linked to questions of power and control. Within the digital environment, power is manifested *among others* through the control digital actors have

[226] See for instance: Michael Gentzel, 'Biased Face Recognition Technology Used by Government: A Problem for Liberal Democracy', *Philosophy & Technology* 34, no. 4 (2021): 1639–63, https://doi.org/10.1007/s13347-021-00478-z.

over data infrastructures, decision-making capacities, data knowledge, etc.[227] From a decisional perspective, this power translates into a capacity to determine how, and under what conditions, algorithmic systems will be imposed; and in many cases, this also entails control over how people will be judged and differentiated against. Undoubtedly, these positions of powers can be (ab)used in ways that maintain and reify existing and systemic forms of social inequality.

When a person, group or entity is in a position of power, they can exercise a greater degree of control over the relationships people stand in, and in doing so, their actions may (inadvertently) give way to novel forms of social inequality.[228] In this context, social equality regulates the responsible use of algorithmic power and the potential algorithmic systems have for interfering in the relational fabric of society. When we apply conceptualisations of social equality onto the algorithmic environment, two general relationships are subject to evaluation. On the one hand, social equality governs the vertical relationship between the decision-maker and the decision-subject. On the other hand, social equality governs the horizontal relationships that are produced and maintained as part of algorithmic decision-making processes.[229]

GOVERNING THE VERTICAL RELATIONSHIP BETWEEN DECISION-MAKER AND DECISION-SUBJECTS –In most cases, the algorithmic decision-maker and decision-subject do not stand in a relation of equality. Instead, the former will have the capacity to exert a certain degree of control over the latter. Given this position of relational inequality, decision-makers should be asked to carefully reflect upon their treatment of decision-subjects: does a given decision-making act have the potential to undermine a decision-subject's standing as a free and equal member of society? In the vertical relationship between decision-makers and decision-subjects, social equality will govern how and under what conditions, algorithmically guided decisions can be imposed. Here, social equality gains a monitoring function. Decision-makers ought to be condemned when they use their powers to dominate or oppress their decision-subjects or, when they expose their decision-subjects to practices that are experienced as degrading, demeaning, inferiorising, stereotyping or

[227] In addition, power can be seen as a function of social, financial, or legal clout. See also Chapter 6. See also: Fisher and Streinz, 'Confronting Data Inequality'.

[228] Vedder and Naudts, 'Accountability for the Use of Algorithms in a Big Data Environment'.

[229] On the relational dimensions of data production, and the egalitarian harms data-driven practices produce see also: Viljoen, 'A Relational Theory of Data Governance', sec. II. Data Relations and Their Social Effects.

stigmatising.[230] The latter may also include the imposition of so-called representational harms onto decision-subjects by the decision-maker.[231]

GOVERNING THE HORIZONTAL RELATIONSHIPS GENERATED THROUGH ALGORITHMS – In analysing their position (of power), a decision-maker should also reflect upon the impact they have on the horizontal relationships decision-subjects stand in. To start with, the (vertical) imposition of representational harms onto (groups of) decision-subjects could have direct bearings for a given person or group's social standing. Gender stereotypes in recommender systems for instance maintain social inequality between men and women.

Yet, people's social standing can be affected simply by virtue of the relational dynamics algorithmic systems impose onto decision-subjects. In general, the production through inferential analytics of population- or group-level knowledge, often an integral component of algorithmically guided decision-making systems, can be described as a process that maps people and their

[230] In this regard, Viljoen rightfully observes how social egalitarian concerns capture or overlap with dignitarian concerns. Viljoen, 631.

[231] The ways in which people are categorised can have an impact on a person's relationship towards themselves: it is better to be considered a good applicant, rather than a bad one, to be trustworthy, rather than non-trustworthy, etc. Furthermore, algorithms could have an inferiorising effect without information concerning the actual classifications used being public: a person's systematic exposure to negative algorithmic outcomes could be sufficient to trigger a negative emotive response vis-à-vis a person's sense of self. In the latter scenario, the person only knows they had an individual loss in a situation where a more beneficial outcome was possible. Another example. Suppose every border control algorithm is built to single out people with a particular, rare eye pattern. People who have this pattern must endure border checks every time they travel. The repeated encounters with security may work alienating and inferiorising. The person in question is a white, heterosexual, cis man. He carries no societal baggage. Yet, if social equality wants to address feelings of inferiorising treatment, then this is certainly a case where the relevance of social equality is triggered. Feelings of inferiority might moreover be compounded within the algorithmic environment. In their research, Lee discovered how algorithmic decisions tend to evoke more negative emotions than human decisions, among others because the evaluation through machines is considered dehumanising. Reference can also be made to other contextual factors that can contribute to a person's sense of self. The opacity of decision-making systems or a lack of transparency may cause a person to feel more badly because they are given no insight as to why they were treated the way they were. See: Min Kyung Lee, 'Understanding Perception of Algorithmic Decisions: Fairness, Trust, and Emotion in Response to Algorithmic Management', *Big Data & Society* 5, no. 1 (1 January 2018): 2053951718756684, https://doi.org/10.1177/2053951718756684. Solon Barocas, Moritz Hardt, and Arvind Narayanan, 'Fairness and Machine Learning', n.d., 32; Reuben Binns, 'Fairness in Machine Learning: Lessons from Political Philosophy', in *Proceedings of Machine Learning Research* (Conference on Fairness, Accountability and Transparency, New York City, 2018), 8–9.

relationship to one another.[232] Likewise, systems that depend upon the categorisation of individuals to attain their goal or to reach a decision are relational. Group members share a relationship of similarity due to the characteristics they (supposedly) have in common. In addition, within each specific decision-making instance, each identified group, and their respective members, stand in relation to one another. For every identified group, there is an out-group.[233]

The main point to be made with regard to algorithms' horizontal and relational dynamics is twofold. First, algorithms operate in relational ways, and each of these operations function as an instance of interference into the relational dynamics within society. Data production, as Viljoen strikingly points out, produces winners and losers.[234] The capacity for algorithmic decision-makers to divide the population into broad, but nonetheless consequential categories, could have great socio-relational repercussions: any relational interference can lower, maintain or increase a person or group's social standing or status, and in doing so, these interferences can come into conflict with socio-relational ideals of equality that people should be treated and recognised as equals in the relationships they hold.[235] Second, as algorithmic decision-makers have the capacity to capture and differentiate between people and groups of people on a variety of grounds, for a variety of purposes, and within a variety of contexts, these relational shifts can be imposed and harm the equal social position of people and groups, irrespective of their current social position.

---

[232] See on this point: Viljoen, 'A Relational Theory of Data Governance'. According to Viljoen, to derive population-level insights regarding the various ways in which data subjects relate to one another, is the core goal of many contemporary data collection practices.

[233] Moreover, this out-group need not be comprised of people who are actually subject to a specific decision. In principle, algorithmic categories cover a wide part of the population. For instance, suppose a government builds an algorithm that will classify people who are unemployed into several categories in relation to the aid they might need for re-entering the job market. The categories used reflect a way of thinking that applies onto every unemployed job seeker within Belgium, and not only those who will actually make use of the algorithm.

[234] Viljoen, 'A Relational Theory of Data Governance', 583.

[235] This might be a particular problem for those decisions that could be considered "hierarchical" in nature. Certain decision-making instances have a clear positive and a clear negative outcome. As observed in an earlier footnote, it is better to be considered a good, rather than a bad applicant, a trustworthy, rather than a non-trustworthy lender. One can imagine that if people discover that certain characteristics are associated with a negative outcome, this might affect their (social) perception of those who possess that particular characteristic. Suppose blue-eyed people are considered as non-trustworthy by financial institutions. Could this have an impact as to how they are perceived by non-blue-eyed people? Of course, within the group of trustworthy applicants, group members may still remain in competition. A financial institution cannot give credit to every applicant. Yet, those who have non-blue-eyes are on equal social standing with one another, even though it could be impossible to treat them equally in terms of the (distributive) outcomes.

For decision-subjects the relationships they are placed in might not be tangible initially. The opacity of decision-making systems, combined with a lack of transparency, might make it impossible to know how one is grouped. Furthermore, because the groups in question lack social saliency, a given negative effect might not be recognised as a group or shared harm: decision-subjects do not necessarily understand their connections with those who are in the same position in relational terms.[236] As a consequence, these social dynamics might take place in ways that are visible or invisible; in ways that take immediate effect, or only arise over time.[237] Yet, despite all of the aforementioned, these social dynamics do exist and they take place on a continuous basis. It therefore does not seem unreasonable to assume that relational similarities and differences could at one point be translated into real and concrete social inequalities.

In this context, the presence of power asymmetries, in combination with the scale at which algorithms can be rolled out, and hence their effects amplified, data-driven differentiation could prompt greater social inequality and disparity.[238] The risk towards horizontal social inequality might be greater where decision-makers have influence over decisions that govern a so-called "realms of opportunity", decisions that affect people's future life choices and paths, and as a consequence, their overall social standing, such as hiring or credit scoring applications.[239]

---

[236] As a consequence of this however, it will be more difficult for groups to come together and fight against a given decision. Similarly, wrongful adverse treatment may be more difficult to detect where less tangible, or seemingly random groups are involved. See also on the latter points: Anton Vedder, 'Medical Data, New Information Technologies and the Need for Normative Principles Other than Privacy Rules', *Law and Medicine. M. Freeman and A. Lewis (Eds.),(Series Current Legal Issues)*, 2000, 441–59.

[237] Liu et al., 'Delayed Impact of Fair Machine Learning'.

[238] Naturally, positions of power may also exacerbate existing the disadvantage experienced by historically disadvantaged groups. Moreover, in relation to historically disadvantaged or protected groups, the above-mentioned social effects can effectuate more swiftly. The use of gender stereotypes in movie recommender systems for instance is such a socio-relational harm. Even if the distributed good in question holds little social importance as a status-good, the use of stereotyped categories could be considered problematic simply because it reinforces a generalised assumption regarding men and women's position in society. The greater a good's social value, and the more salient the group affected is, the greater the societal impact of differential treatment likely will be. Vedder and Naudts, 'Accountability for the Use of Algorithms in a Big Data Environment'.

[239] See also Chapter 7, Section 7.3: Capability Approach. Kathleen Creel and Deborah Hellman, 'The Algorithmic Leviathan: Arbitrariness, Fairness, and Opportunity in Algorithmic Decision Making Systems', in *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency* (FAccT '21: 2021 ACM Conference on Fairness, Accountability, and Transparency, Virtual Event Canada: ACM, 2021), 816–816, https://doi.org/10.1145/3442188.3445942.

Likewise, a given treatment might be experienced as more harmful where these harms have been imposed by an actor who has great social or political clout.

Of course, not every algorithmic decision will have an impact on people's relationships in ways that are worthy of regulation. For instance, the relationships generated by analytical processes can be highly dynamic and evolutive, and hence their socio-relational effect might not be long-lasting. Furthermore, while the data-driven environment might create winners and losers on a continuous basis, the end-result might maintain the status quo. However, one should at least recognise that algorithms pose an inherent risk to social equality.

Social equality encourages regulators to carefully consider to what extent algorithmic decision-makers can be allowed to intervene into the social fabric of society. Likewise, social equality encourages decision-makers to reflect upon their direct relationship with decision-subjects, and the way they can change relational dynamics, and whether these dynamics risk undermining a given person or group's social status or standing.[240]

Algorithmic Distributive Harms

Following a distributive approach towards equality, a society, policy, or regulatory framework will try to identify a given interest that people share (e.g., freedom and self-determination; equal access to employment), a series of (non-)material justice-relevant goods that people should have roughly equal access to in order to safeguard that interest (e.g. equal access to primary goods; equal opportunity when entering the job market). In doing so, a specific, distributive rule will specify how equality in the enjoyment of that interest can be achieved (e.g., the principles of justice; prohibition for employers to discriminate on the basis of certain protected characteristics).

Having become an integral and integrated component of the basic institutions that shape society, algorithmic structures and decision-making acts co-govern the distribution of justice-relevant goods. Distributive views on equality thereby have an intuitive appeal as they can be mapped onto

[240] Of course, there are challenges to this approach. First, See also: Viljoen, 'A Relational Theory of Data Governance', 631. See also the work of: Noble, *Algorithms of Oppression*; D'Ignazio and Klein, *Data Feminism*; O'Neil, *Weapons of Math Destruction*; Crawford, *Atlas of AI: Power, Politics, and the Planetary Costs of Artificial Intelligence*. Julie E. Cohen, *Between Truth and Power: The Legal Constructions of Informational Capitalism* (Oxford, New York: Oxford University Press, 2019). Shoshana Zuboff, *The Age of Surveillance Capitalism: The Fight for the Future at the New Frontier of Power* (London: Profile Books, 2019).

algorithms in their function as a producer of outcomes/distributor of burdens and benefits, and the distributive harms they might generate in the process.[241]

Algorithms are often deployed in domains where they directly decide upon a person's opportunities (e.g., enrolment algorithms) or allocate important benefits or burdens (e.g., welfare allocation). As such, they are typically seen as distributional artefacts that have an immediate impact on people's fundamental interests. In doing so, algorithmic categorisations often map people to outcome classes that take on the form of either a positive or negative outcome (e.g., (not) accepted into higher education/(not) granted welfare). As Binns observes, for many decisions those outcome classes represent a means or a barrier towards the realisation of a fundamental interest.[242]

Distributive conceptualisations of equality could help us in our evaluation of these decision-making instances. For instance, suppose society has defined the right to fair equality of opportunity towards higher education as a justice-relevant good. In order to ensure equal access to fair equality of opportunity, one could argue that enrolment algorithms designed to select and allocate students to university programs should ensure that every student has a fair chance to enter into university. One could argue that the differentiation criteria those enrolment algorithms maintain, should reflect students' individual merit and potential as much as possible, rather than take into account other attributes, such as historical school results.[243]

---

[241] Barocas, Hardt, and Narayanan, 'Fairness and Machine Learning', 32; Binns, 'Fairness in Machine Learning: Lessons from Political Philosophy'.

[242] Binns, 'Fairness in Machine Learning: Lessons from Political Philosophy', 6.

[243] In 2020, the education system had to adapt to the COVID-19 pandemic. One such area that necessitated adaptation was the way in which exams were traditionally organised. Given the impossibility to organise exam sessions attended by a large number of students, different methods of examination had to be found. Education institute International Baccalaureate was among those looking for new evaluation methods. Because the exams had to be cancelled, they devised an algorithm that would predict student grades as an alternative grading mechanism. The attributes used included the coursework scores, grades predicted by teachers and schools' historical IB results. In a note provided by the IB, the following was stated with regard to schools for which there was no historic data: "We are able to award grades for schools for which we have no historic data. In these instances, in place of the historic data, we will use a formula based on the pooled information from every school taking that subject, based on coursework marks and predicted grades." According to the International Baccalaureate their system was the fairest approach. The grading system was met with great disapproval though. While the Norwegian DPA warned against the system's capacity to discriminate on the basis of socio-economic grounds, it also criticised the IBO's failure to sufficiently take into account students' individual merit and potential: "Through the use of 'school context' and 'historical data', the IBO has utilised information concerning how other students at the same school have performed historically to adjust the final grades. This means that the IBO expects students to foresee and accept that their grade is influenced by a factor completely

The influence algorithmic structures and acts have on people's distributive shares often runs deeper. The aforementioned enrolment algorithm might not only affect a person in their access to the justice-relevant good of education but also their access to other justice-relevant goods such as employment, wealth, and income. In this context, algorithmic structures and decision-making acts can be assessed in terms of the overall impact they can have on the desired distribution of justice-relevant goods: do algorithmic decision-making structures and acts, including the socio-relational and socio-economic inequalities they produce, contribute to, or instead interfere with, the just distribution of, or equal access to, those goods? If algorithms interfere with the just or desired distribution of goods, and hence, limit people in their ability to enjoy their fundamental interests, then they are capable of producing an allocative or distributive egalitarian harm.[244] As equalisation is the goal, this harm can take place regardless of the differentiation ground that was used, or the group covered by that ground. [245]

outside their control, which does not have any connection to their demonstrable and individual academic achievements, and that relates to other people. This assumption is unjustified." The Norwegian DPA argued that in doing so the system ran counter to the GDPR's fairness principle, which concerns (*among others)* the reasonable expectations of the people whose data have been collected and processed. In this particular case, the students could not have reasonably expected that their education would be determined on the basis of the criteria used by the algorithm. 'The Norwegian DPA Intends to Order Rectification of IB Grades', Datatilsynet, accessed 18 October 2020, https://www.datatilsynet.no/en/news/2020/the-norwegian-dpa-intends-to-order-rectification-of-ib-grades/.
Datatilsynet (Norwegian Data Protection Authority), 'Advance Notification of Order to Rectify Unfairly Processed and Incorrect Personal Data - International Baccalaureate Organization'. In reference to: European Data Protection Board: Guidelines 2/2019 on the processing of personal data under Article 6(1)(b) GDPR in the context of the provision of online services to data subjects, version 2.0, adopted 8 October 2019, para. 12. Tom Simonite, 'Meet the Secret Algorithm That's Keeping Students Out of College', *Wired*, 2020, https://www.wired.com/story/algorithm-set-students-grades-altered-futures/; Avi Asher-Shapiro, 'Explainer - Exam Grading Algorithms amid Coronavirus: What's the Row About?', *Reuters*, 12 August 2020, sec. Healthcare, https://www.reuters.com/article/global-tech-education-idUSL8N2FD0EI; Han, 'An Algorithm Shouldn't Decide a Student's Future'. Han.
See also International Baccalaureate, 'IB Assessment for May 2020 Examination Session during COVID-19 Pandemic'.

[244] Next to the socio-relational egalitarian harms they might generate. As mentioned in Section 2.1.4, this interest could also be the preservation of socio-relational ideal!

[245] The more accurately defined a given distributional scheme is, the more it can function as a clear guide to decision-makers, and the easier it is to identify distributive harms. Substantive non-discrimination laws for example have clearly defined rules on how certain opportunities, goods or service should be distributed within society.

2.2.2 Algorithmic Equality of Whom?

Arguments have been made that justice and the law should be primarily concerned with the disadvantage faced by social or socially salient groups.[246] Young defines social groups as "a collective of persons differentiated from at least one other group by cultural forms, practices, or way of life."[247] Due to their shared affinity and experience, the members of these groups will associate with one another more than they do with non-group members.[248] Social groups are not fixed, they are fluid and subject to evolution. Social groups stand in relation to other groups and group identification arises following social interactions and encounters.[249] To assess social phenomena like inequality in group-terms does not mean that individuals cannot be the final target of a policy.[250] Young nonetheless notes how "in pursuing individual and institutional goals, [… a] comparison of groups is necessary to judge inequalities unjust because doing so helps show that different individuals have different opportunities and inhibitions related to their structural social positions."[251] The importance of ascribing rights or equality claims to well-defined groups can also be seen in ethical and legal scholarship where some authors have argued that certain differentiation acts should only be of concern when they can be tied to (socially) salient groups.[252]

---

[246] See for instance also: Wachter, 'The Theory of Artificial Immutability', 26–28.

[247] For more information on Young's views on social groups, social difference and identity, as well as the role of the individual within groups; see also: Young, *Inclusion and Democracy*, chaps 3, Social Difference as a Political Resource. See also: Young, *Justice and the Politics of Difference*, chaps 2, Five Faces of Oppression, and in particular the section: 'The Concept of a Social Group'. Young, 'Equality of Whom? Social Groups and Judgments of Injustice', 15.

[248] Young, *Justice and the Politics of Difference*, 43.

[249] Young, 47–48.

[250] Young, 43. For more information on Young's views on social groups, social difference and identity, as well as the role of the individual within groups; see also: Young, *Inclusion and Democracy*, chaps 3, Social Difference as a Political Resource. See also: Young, *Justice and the Politics of Difference*, chaps 2, Five Faces of Oppression, and in particular the section: 'The Concept of a Social Group'. For instance, social groups are not defined by a set of attributes, but through a sense of identity, and even though groups do not exist as independent from individuals, they constitute individuals: a person's being is shaped by their group affinities. This however does not exclude an individual to reject or rise above the groups they find themselves in. Yet, Young does observe that "one first finds a group identity as given, and then takes it up in a certain way." Still, "individual persons, as constituted partly by their group affinities and relations, cannot be unified, themselves are heterogeneous and not necessarily coherent." Young, 45–48. Nor, would justice require the "melting away of differences." Young, Justice and the Politics of Difference, 47.

[251] Young, 'Equality of Whom? Social Groups and Judgments of Injustice', 17. In this regard, Young critiques theories of justice for often being too focused on the individual. See on this point also the views of Nussbaum on the role of structures and the individual demands of justice (Section 8.2.2.). Young, *Justice and the Politics of Difference*, 45.

[252] Kasper Lippert-rasmussen, 'The Badness of Discrimination', *Ethical Theory and Moral Practice* 9, no. 2 (April 2006): 167–85, https://doi.org/10.1007/s10677-006-9014-x; Kasper Lippert-Rasmussen, 'Discrimination: An Intriguing but Underexplored Issue in Ethics and Political Philosophy', *Moral Philosophy and Politics; Berlin* 2, no. 2 (2015): 207–17,

This social group criterion is often complemented with a power criterion. Under this view, justice and the law should address the claims of those who are oppressed and dominated, rather than those in power.

Algorithmic inequalities can be tied to groups that have faced historical disadvantage, or are otherwise socially salient, but in fact *they do not need to be*. When algorithmic decision-makers derive and apply information onto people, the differentiation grounds that have been defined seldom refer to socially salient characteristics or groups. Similarly, the connections they find between people are not always socially salient either. Instead, they are what Young would refer to as aggregates: a "classification of persons according to some attributes."[253] Groups captured by such algorithmic differentiation grounds tend to be non-representational in the sense that they do not necessarily respond to or correspond with more concrete forms of injustice that people have actually seen or experienced in the world.[254]

One reason as to why non-social groups or aggregates would be excluded from justice or the law, is because in our evaluation of the justified or non-justified nature of a given situation, we seldom refer to random attributes.[255] Young believes that "when we criticise social conditions for being

---

http://dx.doi.org.kuleuven.ezproxy.kuleuven.be/10.1515/mopp-2015-0026. In their report on algorithmic decision-making, Castellucia and Le Métayer observe how "existing laws focus on the protection of well-identified social groups (e.g. based on gender, ethnic origin, religion, etc.) but ADS make it possible to discriminate according to many other criteria that could also be considered as unfair in certain situations. Such unfairness could target virtual groups that are not necessarily identified in society (such as, to take a few random examples, left-handed people, people who have learned the same foreign language, or like the same movies). Where should the line be drawn and under what principles?" Castelluccia and Le Metayer, 'Understanding Algorithmic Decision-Making: Opportunities and Challenges', 63.

[253] Young, *Justice and the Politics of Difference*, 43.

[254] The notion of non-representational was initially introduced by Gutwirth and De Hert to describe that "Profiles do not so much aim to represent a current state of affairs but rather to predict future behaviours inferred from past actions." Serge Gutwirth and Paul De Hert, 'Regulating Profiling in a Democratic Constitutional State', in *Profiling the European Citizen: Cross-Disciplinary Perspectives*, ed. Mireille Hildebrandt and Serge Gutwirth (Dordrecht: Springer Netherlands, 2008), 271–302, https://doi.org/10.1007/978-1-4020-6914-7_14; Gloria González Fuster, Serge Gutwirth, and Erika Ellyne, 'Profiling in the European Union':, June 2010, 12; Leese, 'The New Profiling: Algorithms, Black Boxes, and the Failure of Anti-Discriminatory Safeguards in the European Union', 504. They often do not represent actual, real-life social groups, nor do they, as prefaced by Mann and Matzner "make sense as the intersectional combination of identities." Mann and Matzner, 'Challenging Algorithmic Profiling', 5. Mann and Matzner refer to these new forms of differentiation as "emergent discrimination": "both the input and output of algorithmic systems may not have a direct relation to a protected ground but it might still be the case that an algorithmic system systematically disadvantages persons."

[255] I shall make use of the term groups to describe both social groups and algorithmic aggregates.

unjust, we do not choose random attributes of individuals, like their taste in music, or their bone structure, as a basis for grouping them [but] we construct the groups for comparison according to generally recognised social position which we know have broad implications for how people relate to one another [such as] class, race, ethnicity, age, gender [etc.]."[256] Similarly, arguments have been made that seemingly random attributes, such as a person's shoe size, do not affect a person's social standing.[257]

Yet, the introduction of algorithms has obviously changed this dynamic. As seen in Chapter 1, and as further established in the previous section, algorithms have the potential to alter the relations people stand in and the opportunities, burdens and benefits they are given. Furthermore, in doing so, these processes are indiscriminate: anyone can incur socio-relational and distributive harms through their functioning. In other words, even the seemingly random attributes that are used to group people can have broad implications for a person's social position. Indeed, as mentioned before, a society's commitment to equality typically represents its commitment to equalise the enjoyment of certain fundamental interests for all. As Wachter points out, the harms and interests faced by algorithmic and social groups are ultimately the same.[258] Simply put, to address only those inequalities faced by social or socially salient groups, would be to underestimate the (re)structuring power algorithmically guided decision-making systems have as meaning-givers and the way these technological innovations have transformed the way decisions are made. To exclude aggregates from protection is to limit the (a priori) the debate as to what algorithmic justice might require.

Of course, (historically disadvantaged) social groups are still likely to face the greatest harms following the deployment of algorithmically guided decision-making systems. Hence their disadvantage should continue to be most of our concern. Yet, we should not exclude an entire range of algorithmic decision-making processes from entering the realm of either justice or the law. Instead, we should more thoroughly investigate how the law and justice would respond to these claims, and whether through this investigation, we can find common ground through which both social and non-social groups can find protection.

---

[256] Young, 'Equality of Whom? Social Groups and Judgments of Injustice', 15.

[257] Lars Lindblom, 'In Defense of Rawlsian Fair Equality of Opportunity', *Philosophical Papers* 47, no. 2 (4 May 2018): 240, https://doi.org/10.1080/05568641.2018.1445550. In reference to: David Edmonds, *Caste Wars: A Philosophy of Discrimination* (London; New York: Routledge, 2006), http://site.ebrary.com/id/10272708.

[258] Wachter, 'The Theory of Artificial Immutability'.

Finally, an open approach towards the "equality of whom?" question might also better capture the complex nature of group-level inequalities within the digital environment and the harms they might impose onto individuals. Often people do not experience disadvantages due to their possession of a specific, individual trait, such as ethnicity, their physical or mental capabilities, or their sexual orientation. Algorithmic differentiation is rarely single axis, but often complex, multidimensional, and sometimes intersectional in nature.[259]

2.3 Algorithmic Inequality: Related Notions

This Dissertation focuses on the notion of equality as a conceptual lens to evaluate the impact of algorithmically guided decision-making systems. Within algorithmic governance debates, the notion of equality is often used in tandem with three related, but nonetheless distinct notions: algorithmic bias, discrimination, and intersectionality. In this section, the relationship between equality and these notions will be explained.

Bias and Inequality

Bias generally and neutrally refers to a type of predisposition in decisions or judgments. Within decision-making systems, bias can be seen as both a cause and driver towards (unjustified) inequalities. In their seminal 1996 paper "Bias in Computer Systems" Friedman and Nissenbaum distinguish three categories of bias in computer systems that remains apt today: pre-existing, technical, and emergent.[260]

Pre-existing biases refer to the existing or historical attitudes present within society at large, subcultures, institutions or the personal attitudes of individuals that are placed, either consciously or unconsciously, explicit, or implicit, into the computer system.[261] Technical biases stem from technical constrains or considerations. Such predispositions can arise due to the designers' attempt to technically constrain human constructs that are not fit for such a purpose, e.g., "to

[259] See also below, Chapter 2, Section 2.2.3. "Equality and Intersectional Differentiation."

[260] Friedman and Nissenbaum, 'Bias in Computer Systems'. In recent years, a lot more research into computer bias has emerged. A 2019 study by Mehrabi and others for example distinguished 23 possible sources of data bias: Mehrabi et al., 'A Survey on Bias and Fairness in Machine Learning'. See also: Eirini Ntoutsi et al., 'Bias in Data-Driven Artificial Intelligence Systems—An Introductory Survey', *WIREs Data Mining and Knowledge Discovery* 10, no. 3 (2020): e1356, https://doi.org/10.1002/widm.1356.

[261] Friedman and Nissenbaum, 'Bias in Computer Systems', 333.

quantify the qualitative, discretize the continuous or formalize the non-formal."[262] Legal expert systems for instance try to automate certain parts of the legal process but often fail to capture the many nuances of legal reasoning by assuming that the law can be caught and translated in an unambiguous manner. Likewise, formal fairness metrics designed to ensure equal outcomes within decision-making systems might oversimplify substantive egalitarian objectives.

Finally, Friedman and Nissenbaum identify emergent bias. Emergent bias occurs due to a system's interaction with a real environment and real users. It takes place after the design of a system has been completed and is the result of shifting or changing societal knowledge, population, or cultural values.[263] Two general types of emergent bias can be further identified: bias due to new societal knowledge and bias due to a mismatch between users and system design. The first type of bias relates to information that is discovered after a system has been developed. Because this information could not be integrated during the design, it is disregarded within the system's future functioning. For instance, an antiquated welfare allocation algorithm might fail to adequately capture new social vulnerabilities that have emerged due to a growth in population If left unaddressed, the programme will favour the old, rather than updated population statistics, and leave those with genuine and current needs in a bad position. A bias from mismatch will emerge due to differences between a system's actual end-users and the end-users assumed during the design stage.[264] For instance, educational software designed to reward children's hard-skills might not correspond with the educational needs of parents and teachers as these technologies might discourage children from developing their soft-skills.[265]

Most decision-making processes are biased in the sense that they will favour a particular outcome over the other. Bias becomes problematic when it results in treatment that is undesirable. This may be the case when the outcome favoured does not correspond to the needs of the end-users, or when it generates unjustifiable inequalities.[266] If we want to identify and rid algorithmic

---

[262] Friedman and Nissenbaum, 334.

[263] Friedman and Nissenbaum, 335.

[264] Friedman and Nissenbaum's example is slightly different: "educational software to teach mathematics concepts is embedded in a game situation that rewards individualistic and competitive strategies, but is used by students with a cultural background that largely eschews competition and instead promotes cooperative endeavors."

[265] Mehrabi et al., 'A Survey on Bias and Fairness in Machine Learning'.

[266] Especially where algorithms are built to learn from their own past experiences, bias can generate so-called feedback loops whereby the gap between the groups differentiated becomes bigger. In so doing, algorithms can introduce greater inequality on a social or relational and socio-economic or distributive level. See for instance: Ensign et al., 'Runaway Feedback Loops in Predictive Policing'. In the case of bias vis-à-vis historically disadvantaged group, the problem is often not only that bias replicates disadvantage, but that bias reinforces and *amplifies* disadvantage!

decision-making systems from bias because they build upon, perpetuate, reflect, or introduce unjustified inequalities, regulators and decision-makers should be guided by a concrete conceptualisation of equality. The latter will determine what biases represent a genuine egalitarian harm. In this regard, Balayn and Gürses rightfully remark how "policymakers should engage with and learn from prior work on eliminating discrimination and inequalities as part of identifying and tackling AI's impact on inequalities."[267] Put differently, policy makers should first identify, articulate, and make explicit what inequalities they consider unjustified before they encourage algorithmic decision-makers to design their system in such a way that they preclude biases that are likely to replicate or reinforce these harms.[268]

Equality and (Non-)Discrimination

From a legal perspective, the concept of equality is often used in tandem with the concept of non-discrimination. While equality and (non-)discrimination are strongly related notions, they are nonetheless conceptually distinct.[269] Discrimination generally refers to the adverse or disadvantageous treatment imposed upon, or adverse impact experienced by, a person or group of persons. For instance, direct racial discrimination concerns the situation where a person is treated less favourably than another on grounds of race or ethnic origin. Indirect racial discrimination occurs where a neutral rule places persons of a certain racial or ethnic origin at a particular disadvantage as opposed to others. Both types of adversity can come into conflict one or more ideals of equality holds, and in doing so, impose egalitarian harm onto people.

Suppose a government welfare policy would deny social security benefits to people of colour. For people of colour, this treatment represents a moral egalitarian harm (their interests seemingly do not matter equally), a socio-relational egalitarian harm (they are treated as being of lesser social standing) and a distributional egalitarian harm (they are deprived of a benefit everyone should have equal access to).[270] Discrimination is prohibited due to this conflict with the principle of

[267] Agathe Balayn and Seda Gurses, 'Beyond Debiasing: Regulating AI and Its Inequalities.' (EDRi, 2021), 14.

[268] One could also grant decision-makers a certain degree of discretion. They could advance their own conceptualisation of equality but make explicit their design choices through mechanisms of accountability.

[269] For a more thorough discussion on the relationship between non-discrimination and equality, see also Johanna Croon-Gestefeld, *Reconceptualising European Equality Law: A Comparative Institutional Analysis*, 1st ed., Modern Studies in European Law (Oxford: Hart Publishing, 2017), chap. 2, at page 46, https://doi.org/10.5040/9781509909674.

[270] In a similar vein, Shin has argued that the "best moral explanation for the legal concept of discrimination is that it embodies a collection of approaches that together express our society's commitment to identify, disavow, and disallow in our institutional practices the categories of actions that tend to reinforce or resonate with historic and persistent

equality.[271] For the remainder of this dissertation, the relationship between equality and non-discrimination will be understood thusly.[272]

Still, laws and policies may differ in the way they conceptualise equality and discrimination as notions.[273] For example, discrimination can be defined as the adverse treatment of *socially salient*

---

patterns of unjust inequalities. Patrick S. Shin, 'Is There a Unitary Concept of Discrimination?', in *Philosophical Foundations of Discrimination Law* (Oxford: Oxford University Press, 2013), https://doi.org/10.1093/acprof:oso/9780199664313.003.0009. See also: Dagmar Schiek, 'From European Union Non-Discrimination Law towards Multidimensional Equality Law for Europe', in *European Union Non-Discrimination Law: Comparative Perspectives on Multidimensional Equality Law* (Routledge-Cavendish Taylor & Francis Group, 2008), 3–27, https://doi.org/10.4324/9780203892626.

[271] In this regard, Muir observes how legal scholars tend to view equality as an overarching policy objective and discrimination as (one) symptom of society's failure to achieve that goal. Elise Muir, *EU Equality Law: The First Fundamental Rights Policy of the EU*, Oxford Studies in European Law (Oxford: Oxford University Press, 2018), 21, https://doi.org/10.1093/oso/9780198814665.001.0001. This relationship between equality and discrimination is not shared by everyone. See for instance: Elisa Holmes, 'Anti-Discrimination Rights without Equality', *The Modern Law Review* 68, no. 2 (2005): 175–94. For a discussion of Holmes, reference can also be made to the work of Croon-Gestefeld who analyses and critiques the conclusions of Holmes. Croon-Gestefeld, *Reconceptualising European Equality Law: A Comparative Institutional Analysis*, chap. 2, at pages 40-44. Other authors approach discrimination as a moral wrong and equality as a principle of justice. Under this view, certain forms of discrimination can be justified in light of justice. Under this view, Vallentyne observes "all that matters for justice is that people get their fair share of resources (including social resources, such as opportunities) and this is compatible with being subject to discrimination." However, in most of these discussions, one could argue that discrimination as a moral wrong constitutes violation of equality as a relational ideal. Kasper Lippert-Rasmussen, *Discrimination*, ed. Serena Olsaretti, vol. 1 (Oxford University Press, 2018), https://doi.org/10.1093/oxfordhb/9780199645121.013.20. Lippert-Rasmussen, 'Discrimination', 2015. Peter Vallentyne, 'Discrimination and Rights', in *The Routledge Handbook of the Ethics of Discrimination*, ed. Kasper Lippert-Rasmussen, 1st ed. (New York : Routledge, 2017. | Series: Routledge handbooks in applied ethics: Routledge, 2017), 132, https://doi.org/10.4324/9781315681634-13. See also the many discussions in: Kasper Lipper-Rasmussen, ed., *The Routledge Handbook of the Ethics of Discrimination* (London and New York: Routledge, 2018).

[272] See also: Elise Muir, *EU Equality Law: The First Fundamental Rights Policy of the EU*, Oxford Studies in European Law (Oxford: Oxford University Press, 2018), 21, https://doi.org/10.1093/oso/9780198814665.001.0001. In reference to: Takis Tridimas, *The General Principles of EU Law*, 2nd ed. (New York ; Oxford: Oxford University Press, 2006), 64. Croon-Gestefeld, *Reconceptualising European Equality Law: A Comparative Institutional Analysis*, 46. For instance, the Council of Europe has clarified that non-discrimination and equality should be seen as notions that are closely intertwined. Council of Europe, Explanatory Report to Protocol No. 12 to the Convention for the Protection of Human Rights and Fundamental Freedoms (4 November 2000), para 15. In practice however, the distinction between equality and non-discrimination is not always clear, and both terms tend to be used interchangeably. See also: Muir, *EU Equality Law: The First Fundamental Rights Policy of the EU*, 2018, 21. In reference to: Tridimas, *The General Principles of EU Law*, 64.

[273] Some authors try to formulate the philosophical underpinnings of non-discrimination law. See for instance: Hellman and Moreau, *Philosophical Foundations of Discrimination Law*. Others have set out a more ambitious goal to define a

groups. Under this view, seemingly random groups (or aggregates) would be excluded from the protection of non-discrimination laws.[274] Similarly, law makers may have reason to narrow down the scope of the law. They can prioritise certain dimensions of equality or choose to prohibit certain forms of differential treatment over others. Certain forms of differentiation may therefore constitute an egalitarian harm, but not necessarily a legally prohibited act of discrimination.

The goal of equality may not always be achieved through non-discrimination laws alone.[275] Like equality itself, the problematic nature of discrimination might also be explained alongside other notions, such as dignity and respect.[276] Other values may therefore ground the regulation of differential treatment. The fight against racism can be motivated from both a socio-relational egalitarian and dignitarian perspective.

Following the above-mentioned, the notion algorithmic discrimination will refer to legally problematic forms of algorithmic differentiation. Algorithmic (in)equality is a broader notion, which more generally considers the social position people find themselves in, whether it be in the relationships they hold, or in the distributive shares they have been given. Whereas algorithmic discrimination leads to unjustified inequalities, algorithmic inequalities are not always the result of concrete (algorithmic) discriminatory actions or behaviours.

Equality and Intersectional Differentiation

Intersectional discrimination has been defined as discriminatory treatment based on more than one ground, whereby "either the specific contribution of any one of these grounds is indiscernible or the full extent of discrimination is only recognizable by acknowledging the combination of two

---

theory of discrimination law, to define the core issues non-discrimination as a legal domain should be concerned with. See for example: Solanke, *Discrimination as Stigma*; Khaitan, *A Theory of Discrimination Law*.

[274] See for instance: Lippert-Rasmussen, 'Discrimination', 2015; Lippert-rasmussen, 'The Badness of Discrimination'; Kasper Lippert-Rasmussen, *Born Free and Equal?: A Philosophical Inquiry Into the Nature of Discrimination* (OUP USA, 2014). Khaitan, *A Theory of Discrimination Law*.

[275] Muir, *EU Equality Law: The First Fundamental Rights Policy of the EU*, 2018, 21.

[276] For instance, Hellman has provided an account under which differential treatment (or discrimination) is a wrong when it is demeaning, which she defines as an "expression of denigration by an actor with power to lower their actual social status." To demean a person is morally problematic because it fails to treat people as equals. Deborah Hellman, 'Discrimination and Social Meaning', in *The Routledge Handbook of the Ethics of Discrimination* (Routledge, 2017), 102. See also: Erin Beeghly, 'Discrimination & Disrespect', in *Routledge Handbook to the Ethics of Discrimination*, ed. Kasper Lippert-Rasmussen (New York, NY: Routledge, 2017), 83–96.

or more grounds."[277] For instance a black woman may be discriminated against, not necessarily because she is black, or because she is a women, but due to her identity as a black women.[278] In her account on intersectional discrimination, Atrey notes how "intersectionality rejects the understanding of discrimination as a function of a single categorial axis and emphasizes the need to recognize discrimination resulting from the intersections of multiple axes of race, caste, religion, sex, gender, disability, age, sexual orientation etc."[279]

---

[277] Kimberlé Crenshaw introduced the notion of intersectional discrimination in order to describe the discriminatory experiences of black women: *"Black women can experience discrimination in ways that are both similar to and different from those experienced by white women and Black men. Black women sometimes experience discrimination in ways similar to white women's experiences; sometimes they share very similar experiences with Black men. Yet often they experience double-discrimination-the combined effects if practices which discriminate on the basis of race, and on the basis of sex. And sometimes, they experience discrimination as Black women-not the sum of race and sex discrimination, but as Black women."* Kimberle Crenshaw, 'Demarginalizing the Intersection of Race and Sex: A Black Feminist Critique of Antidiscrimination Doctrine, Feminist Theory and Antiracist Politics', *University of Chicago Legal Forum* 1989, no. 1 (1989): 149. Crenshaw herself noted how intersectionality was not to be seen as a "new, totalizing theory of identity", but instead built upon a history of black feminist scholarship. Kimberle Crenshaw, 'Mapping the Margins: Intersectionality, Identity Politics, and Violence against Women of Color', *Stanford Law Review* 43, no. 6 (1991): 1244, https://doi.org/10.2307/1229039. See also: Dagmar Schiek, 'On Uses, Mis-Uses and Non-Uses of Intersectionality before the Court of Justice (EU)', *International Journal of Discrimination and the Law* 18, no. 2–3 (1 June 2018): 83, https://doi.org/10.1177/1358229118799232. In reference to: Timo Makkonen, 'Multiple, Compound and Intersectional Discrimination: Bringing the Experiences of the Most Marginalized to the Fore.' (Turku: Åbo Akademi University. Institute for Human Rights, 2002), 9–14; Susanne Burri and Dagmar Schiek, 'Opportunities for Legal Responses to Intersectional Gender Discrimination?', 2009, 4–6. For a more in-depth legal exploration of intersectionality as a form of discrimination and theory, see also: Shreya Atrey and Peter Dunne, eds., *Intersectionality and Human Rights Law* (Hart Publishing, 2020), https://doi.org/10.5040/9781509935321; Shreya Atrey, *Intersectional Discrimination*, *Intersectional Discrimination* (Oxford University Press), accessed 8 July 2020, http://www.oxfordscholarship.com/view/10.1093/oso/9780198848950.001.0001/oso-9780198848950; Crenshaw, 'Demarginalizing the Intersection of Race and Sex: A Black Feminist Critique of Antidiscrimination Doctrine, Feminist Theory and Antiracist Politics'; Crenshaw, 'Mapping the Margins'. For intersectionality as a social theory, reference can also be made to the work of Patricia Hill Collins: Patricia Hill Collins, *Intersectionality as Critical Social Theory*, *Intersectionality as Critical Social Theory* (Duke University Press, 2019), https://www-degruyter-com.kuleuven.e-bronnen.be/document/doi/10.1515/9781478007098/html; Patricia Hill Collins, 'Intersectionality's Definitional Dilemmas', *Annual Review of Sociology* 41, no. 1 (14 August 2015): 1–20, https://doi.org/10.1146/annurev-soc-073014-112142.

[278] For an in-depth exploration of various types of discrimination and how they are dealt with in the law, reference can be made to: Atrey, *Intersectional Discrimination*, chap. 3. See also: Raphaele Xenidis, 'Multiple Discrimination in EU Anti-Discrimination Law: Towards Redressing Complex Inequality?', *EU Anti-Discrimination Law Beyond Gender*, 15 November 2018, 46–48, https://www.research.ed.ac.uk/en/publications/multiple-discrimination-in-eu-anti-discrimination-law-towards-red.

[279] Atrey, *Intersectional Discrimination*, 2019, 77. That is not to say that single grounds or attributes cannot have an importance. For example, in describing the relationship between intersectionality and discrimination laws, Atrey notes

HOW ALGORITHMS IMPOSE INTERSECTIONAL DISCRIMINATION - The aforementioned study of Buolamwini and Gebru of gender disparities in commercial gender classification systems is a good example of algorithmic intersectional discrimination.[280] They found that darker-skinned females were the most misclassified group. In their analysis of the Dutch child-benefit algorithms, Amnesty International observed how "different layers of inequality were created through discrimination based on race and ethnicity in the use of the risk classification model, combined with the policies and practices that more often affected people with a lower economic status, who often belonged to an ethnic minority." They conclude: "the childcare benefits scandal must therefore be understood through the lens of intersectional discrimination."[281] Under one discriminatory model, the Austrian AMS employment agency would deduct points to women, disabled people and people over 30.[282] Individuals with the lowest scores ran the risk of being assigned to the least eligible group, and receive less support resources as a consequence.[283] While the system was socially desirable in the sense that it would produce beneficial support outcomes to those individuals who belonged to one or a few disadvantaged groups, individuals who were associated with many data entries belonging to disadvantaged groups ran the risk of being classified the lowest.[284] Through the system's compound effects, the most detrimental outcomes could be experienced by those at the intersection, such as disabled women over thirty years old.[285]

---

how: "Far from splintering identities into unrecognizable and unusable categories, the legal construct of grounds in discrimination law provides a site for thinking about individual and specific instances of intersectional discrimination within a wider context of grounds and groups and thus as a whole." Atrey, *Intersectional Discrimination*, 61.

[280] Buolamwini and Gebru, 'Gender Shades'.

[281] 'Xenophobic Machines: Discrimination through Unregulated Use of Algorithms in the Dutch Childcare Benefits Scandal.' (Amnesty International, October 2021), 23, https://www.amnesty.nl/content/uploads/2021/10/20211014_FINAL_Xenophobic-Machines.pdf?x42580.

[282] Lopez, 'Reinforcing Intersectional Inequality via the AMS Algorithm in Austria'; Kayser-Bril, 'Austria's Employment Agency Rolls out Discriminatory Algorithm, Sees No Problem'. Lopez, 'Bias Does Not Equal Bias'.

[283] Lopez, 'Bias Does Not Equal Bias', 9; Allhutter et al., 'Algorithmic Profiling of Job Seekers in Austria'.

[284] Lopez, 'Bias Does Not Equal Bias', 10.

[285] Even though people might have been deducted points on an additive, rather than a strictly intersectional basis, as observed by Lopez: "[…] If, due to an intersectional convergence of various "point deductions" (the sensitive data entries being age, gender, disability status, childcare responsibilities, nationality) corresponding to one's intersectional positionality, a person is outsourced to group C and does not receive certain support resources for this reason alone, then this is at the very heart of the concept of intersectional discrimination: only the intersection of several axes of vulnerability and therefore only the combination of several "point deductions" leads to a placement in the disadvantaged group C, whereas the existence of "only one" disadvantaged feature implies a placement in the most eligible group B." Lopez, 9–10. Lopez, 9; Lopez, 'Reinforcing Intersectional Inequality via the AMS Algorithm in Austria'; Ben Wagner et al., 'Der AMS-Algorithmus.', *zeitschrift für kritik & recht gesellschaft* 2020, no. 2 (2020): 191–202, https://doi.org/10.33196/juridikum202002019101.

CAN ALGORITHMS IMPOSE INTERSECTIONAL DIFFERENTIATION? - Beyond intersectional discrimination based on sensitive characteristics, such as gender and ethnicity, algorithmic categories may more generally appear intersectional in nature. Indeed, there is a conceptual kinship between the intersection metaphor and the complex ways in which algorithms differentiate.[286]

Assume a law enforcement agency wishes to predict radicalisation among youths. An analysis of social media data and a commissioned socio-psychological study indicates that there will be a higher risk towards right-wing extremism among well-educated twenty-somethings that frequent Reddit user-boards in the category "videogames." Their Internet traffic will be subject to surveillance measures and their Internet conversations will be monitored. Any one of these attributes in isolation is insufficient to introduce such a measure. However, if a person ticks all three boxes, i.e., be a (1) well-educated twenty-something, 2) Reddit user with 3) an interest in video games), they will be deemed as a risk towards national security. This profile serves as a complex differentiation ground that is not reducible to a single attribute. It appears to be intersectional: it is this combined identity that makes them subject to surveillance, and not any single ground on its own.

Of course, the experiences of black women, and people living in low-income areas, should not be equated with the experiences of twenty-something Reddit users. Following Collins (and Carastathis), the "intersection" metaphor should not be used as an "ahistorical, 'universal' construct applied to all sorts of power relationships."[287] Whereas algorithmic categories may appear intersectional in nature or appearance, profiles do not necessarily reflect an "oppressive hierarchy", nor do they target group that are oppressed in multiple ways.[288] Crenshaw's account on intersectionality emphasised the struggles and lived experiences of those who carry multiple burdens.[289] Intersectionality aims to bring "the experiences of the most marginalised to the fore."[290] Lopez rightfully saw how the AMS algorithm functioned in a way that directly went to the heart of the intersectional critique: whereas the system took care of individuals that were "a little"

---

[286] Note that in this particular case, because I do not longer talk about profiles characterised by protected traits, I refer back to the notion "differentiation", rather than "discrimination"

[287] Anna Carastathis, 'Basements and Intersections', *Hypatia* 28, no. 4 (2013): 713. Citing Patricia Hill Collins: Patricia Hill Collins, *Fighting Words: Black Women and the Search for Justice.* (Minneapolis: University of Minnesota Press, 1998), 129, https://www.upress.umn.edu/book-division/books/fighting-words.

[288] Carastathis, 'Basements and Intersections', 711, 712 and 713.

[289] In order to illustrate this point, Crenshaw introduces the "basement" metaphor. Crenshaw, 'Demarginalizing the Intersection of Race and Sex: A Black Feminist Critique of Antidiscrimination Doctrine, Feminist Theory and Antiracist Politics', 151–52.

[290] Makkonen, 'Multiple, Compound and Intersectional Discrimination: Bringing the Experiences of the Most Marginalized to the Fore.'

disadvantaged, it did so in a way that was detrimental to those multiply burdened.[291] Twenty year old Reddit users interested in video games may be unfairly treated and they might experience egalitarian harms, but they are not a socially oppressed group.

Still, intersectionality and the intersection metaphor lay bare the difficulties and challenges involved in both the evaluation and regulation of complex forms of algorithmic differentiation.[292] From a governance perspective, intersectional scholarship has clearly demonstrated how a regulatory paradigm that focuses on distinct and unidimensional groups will fail identify to identify more complex egalitarian harms experienced by those who found themselves multiply burdened.[293]

2.4 The Evaluative Advantage of Equality

Decision-subjects, regulators, as well as the population at large, have a vested interest knowing how people will be judged, treated and differentiated: what attributes or characteristics, behaviours or preferences made a difference within a given context?[294] To obtain access to the innards of a decision-making system, including the underlying logic, design assumptions, the

[291] Lopez, 'Bias Does Not Equal Bias', 10.

[292] Given the opacity of algorithmic decision-making procedures, it may not always be easy to distinguish whether multi-attribute and complex algorithmic profiles are indeed "intersectional" in nature, whether they constitute single-axis (only one identity is part of the discrimination inquiry) or whether they are a multi-axis (multiple identities are part of the discrimination inquiry), but non-intersectional form of discrimination. A complex algorithmic category may in fact turn out to be a proxy for one single, protected ground, and constitute a more simple and straightforward form of indirect discrimination as described under the previous section. For instance, in the case of the Reddit-users, the profile might never have been a proxy for radicalisation, but instead for the protected characteristic political ideology or religious belief. In addition, in case multiple data entries are involved, the concrete bases for differentiation are more difficult to locate. How do attributes interact with one another in shaping a given algorithmic outcome? How, and to what extent, can an algorithmic category burden those affected? All these questions appear relevant to the evaluation of algorithmic decision-making systems, but they are not easily addressed.

[293] See also: Shreya Atrey, 'Beyond Universality : An Intersectional Justification of Human Rights', in *Intersectionality and Human Rights Law*, ed. Shreya Atrey and Peter Dunne, 1st ed. (Oxford: Hart Publishing, 2020), 36, http://www.bloomsburycollections.com/book/intersectionality-and-human-rights-law/ch1-beyond-universality/.
That does not mean non-discrimination law is incapable of addressing intersectional discrimination. Rather, the approach must be recalibrated to fit the difficulties faced by those multiply burdened. See also: See also Atrey, *Intersectional Discrimination*, 60–63.

[294] See on the "differences that make a difference": M. Hildebrandt, 'Discrimination, Data-Driven AI Systems and Practical Reason', *European Data Protection Law Review* 7, no. 3 (2021): 358–66, https://doi.org/10.21552/edpl/2021/3/4.

differences that made a difference, increases the degree of control, agency and democratic oversight external agents can exercise.[295]The attributes, behaviours or preferences that make up a decision-ground form an integral part of the explanation process. It allows decision-subjects to assess what information a decision-maker deemed relevant, and perhaps, what they could have done otherwise to obtain a different outcome.[296] Decision-subjects could verify whether the information used or ascribed to them was accurate or reasonable, and if not, contest the decision, or adduce additional information that the system could or did not (yet) take into account, but of which it is believed it should be of relevance to the decision-maker.[297]

Knowing what information has been considered relevant for a specific decision can help a larger public assess to what extent a decision-making process corresponds to their sense of justice. For example, the Ofqual A-level algorithm, which was set to predict students' grades as students could not take individuals exams due to schools' closure during COVID-period was met with public outcry.[298] One of the factors that Ofqual used as a mechanism for prediction was the historical grade distribution, i.e. the grades other pupils had received, over the previous years.[299] The use of historical school grades as a means to determine the future of pupils was deemed unjust. On the

---

[295] Hildebrandt, in reference to Bateson, observes the following: "Bateson noted that when we perceive our environment we can discern an infinite number of differences between things, and to survive and flourish we need to discern the *differences that make a difference* (to us). If we get this wrong, as a person, as a community, as humanity, we may not survive and will not flourish. In fact, all living beings are in the process of discerning the difference that makes a difference in their environment, as this is what affords organisms to act on relevant differences." In this dissertation I argue that in order to discern those differences that make a difference, we first need to establish the fundamental interests equality is said to safeguard. M. Hildebrandt, 'Discrimination, Data-Driven AI Systems and Practical Reason', *European Data Protection Law Review* 7, no. 3 (2021): 358–66, https://doi.org/10.21552/edpl/2021/3/4. Daniel Foxman and Gregory Bateson, 'Steps to an Ecology of Mind', *The Western Political Quarterly* 26, no. 2 (June 1973): 345, https://doi.org/10.2307/446833.

[296] For instance, under a counterfactual explanation, a decision-subject would be told why a given negative outcome was received, and what would need to change in order to obtain a more positive outcome. In a hiring algorithm, the explanation could be that a decision-subject had too little experience, and had they had one year of seniority at their previous position, they would have been able to apply. Atoosa Kasirzadeh and Andrew Smart, 'The Use and Misuse of Counterfactuals in Ethical Machine Learning', in *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency* (FAccT '21: 2021 ACM Conference on Fairness, Accountability, and Transparency, Virtual Event Canada: ACM, 2021), 228–36, https://doi.org/10.1145/3442188.3445886; Matt J Kusner et al., 'Counterfactual Fairness', n.d., 11; Sandra Wachter, Brent Mittelstadt, and Chris Russell, 'Counterfactual Explanations without Opening the Black Box: Automated Decisions and the GDPR', 2017.

[297] Wachter and Mittelstadt, 'A Right to Reasonable Inferences'.

[298] Hern, 'Ofqual's A-Level Algorithm'; Stewart, Weale, and Proctor, 'Ofqual Chief to Face MPs over Exams Fiasco and Botched Algorithm Grading'. BBC News, 'A-Levels and GCSEs'.

[299] Hern, 'Ofqual's A-Level Algorithm'.

one hand, the use of this criterion was classist as in practice it placed private schools at an advantage compared to state school equivalents.[300] On the other hand, the criterion downplayed the individual merit and potential of pupils. On both points, our evaluation of the attribute in question does not happen in a vacuum. It is shaped by the specific conceptualisation of equality we have in mind, which allows us to see how a specific egalitarian harm can manifest itself through the use particular attribute or differentiation ground. Hence, the articulation of equality allows us to pinpoint the exact injustice involved in the use of certain criteria. In this case, the injustice results from the use of historical school grades which reinforces socio-relational inequality in favour of the elite in an area where we believe a merit-based principle focused on students' individual abilities should govern the allocation of opportunities in education.[301]

We turn to a higher-order notion of equality to guide our construction and evaluation of a decision-making process. [302] Likewise, we can use conceptualisations of equality to identify and criticise the social manifestation of (new forms of) injustice.[303] For example, this is the case where differentiating criteria that appear reasonable in isolation turn out to reinforce the historical disadvantage faced by social groups, or where we believe a system might risk the creation of novel socio-relational and distributional inequalities.[304] Once conceptualised, equality allows decision-

---

[300] A-Levels and GCSEs: How Did the Exam Algorithm Work?', *BBC News*, 20 August 2020, sec. Explainers, https://www.bbc.com/news/explainers-53807730.

[301] This is a crude example: systems that focus solely on merit and ability as isolated traits without considering the wider social conditions under which people can develop their abilities also tend to reinforce social inequality. See for example: Amartya Sen, 'One. Merit and Justice', in *One. Merit and Justice* (Princeton University Press, 2018), 5–16, https://doi.org/10.1515/9780691190334-003; Fred Feldman and Brad Skow, 'Desert', in *The Stanford Encyclopedia of Philosophy*, ed. Edward N. Zalta, Winter 2020 (Metaphysics Research Lab, Stanford University, 2020), https://plato.stanford.edu/archives/win2020/entries/desert/; Richard Arneson, 'Equality of Opportunity', in *The Stanford Encyclopedia of Philosophy*, ed. Edward N. Zalta, Summer 2015 (Metaphysics Research Lab, Stanford University, 2015), https://plato.stanford.edu/archives/sum2015/entries/equal-opportunity/.

[302] Still, equality is but one principle or value however that is at stake within algorithmically guided decision-making systems. For example, algorithms can be critiqued for their predictive properties or their interference into people's choice. As shall be argued throughout this dissertation however, in many cases, inequality is problematic because it interferes with people's dignity or autonomy. Hence all these values are interwoven to establish a totality of governing conditions that algorithmic decisions should adhere to.

[303] In this regard, algorithmic systems might have a delayed impact over time. See also: Liu et al., 'Delayed Impact of Fair Machine Learning'.

[304] For example, a gig-economy delivery service may build a reward system whereby drivers are automatically allocated a higher bonus when they clock in during the night. Workers are classified as deserving a reward when their work schedule takes place at night. This differentiation ground has been chosen by the developer as an incentive for people to work during less-desirable hours of the day. While at first sight this differentiation ground could be rationalised, it could place women at a disproportionate disadvantage. In societies where women are still predominantly assigned to

makers (ex-ante) and evaluators (ex-post) to evaluate the validity of a given differentiation ground.[305]

The evaluation as to whether a decision was fair or unfair, justified, or unjustified cannot always be resolved by looking at the concrete attributes used within a system. Systems might be (purposefully) opaque as to the criteria that matter in a decision-making process.[306] Information once thought relevant for a decision, might lose its relevancy as the result of a continuous learning process. Users' preferences and behaviour for example are subject to change. Hence, it might not always be possible to trace back the reasoning underlying a given treatment to a *specific* behaviour or preference. Likewise, it can be unclear which groups and individuals were targeted by a given decision. As aforementioned, algorithmic decision-making systems can impose egalitarian harms onto concrete, social or tangible groups, as well as non-representational aggregates. In addition, due to the dynamic nature of certain algorithmically guided decision-making processes, the same egalitarian harms might be experienced by different groups at distinct points in time.

By making use of equality as a normative lens, the focus shifts towards the egalitarian disadvantage or harm a system is likely to impose.[307] Algorithms present an injustice where they do not treat people with equal concern and respect; distort the relationships people they stand in; or fail to equalise a fundamental interest through the allocation of opportunities, burdens and risks. By giving concrete shape to these abstract dimensions, the notion of equality can present direct norms for action that can inform decision-makers and regulators as to the conditions under

---

traditional roles in personal relations and the division of labour, such a system may be indirectly disadvantageous towards women. Child caring duties may preclude women from working night shifts. Likewise, general feelings of unsafety associated with city night life might make it more uncomfortable for them to comfortably take deliveries at night. This example is based upon Fletcher's account in: Fletcher, 'Gender Inequality in the Gig Economy', *Autonomy* (blog), June 2021, https://autonomy.work/portfolio/fletcher-gig-economy/.

[305] There is no rule that can universally say when the use of a certain piece of information *in itself* would be wrong without considering the context in which a decision was made. A person's height might not be relevant when a football organisation wants to hire a new IT consultant, but it can be when they are looking for a new goalkeeper.

[306] Burrell, 'How the Machine 'thinks': Understanding Opacity in Machine Learning Algorithms'.

[307] In doing so, equality may be used to bypass the puzzling question that some might claim, lies at the heart of the discussion, namely, to look for those differences that can serve as a justified basis for differentiation. In certain cases, the risk of a disadvantage might be as high that the deployment of the algorithm should be reconsidered altogether.See in this regard also the work of Annette Zimmermann who discusses the question whether certain AI technologies should be built to begin with. Zimmermann, 'The Power of Choosing Not to Build'. In a similar vein, the need to attach certain "labels" or "classes" to people has been criticised, see for instance: Miceli et al., 'Documenting Computer Vision Datasets'. Crawford and Paglen, 'Excavating AI: The Politics of Images in Machine Learning Training Sets'. Crawford, *Atlas of AI: Power, Politics, and the Planetary Costs of Artificial Intelligence*.

which such harms are more likely to occur. [308] Moreover, as I shall hope to demonstrate throughout this dissertation, equality can be used in a more holistic manner in order to guide and evaluate choices made during a system's design, development, or deployment.[309] Indeed, equality can act as a guide throughout the entire algorithmic chain by informing decision-makers what inequalities they should avoid, when they risk imposing such inequalities, and how to design their systems in a way that can avoid unjust inequality.[310]

2.5 Conclusion

In her relational theory of data governance, Viljoen notes how: "datafication may be wrong not only because it manipulates people, but also because the social effects it produces or materializes violate standards of equality. As an economic process, datafication may lead to unfair wealth inequality that violates distributive ideals of justice. As a social process, datafication may reproduce and amplify forms of social hierarchy that violate relational standards of justice."[311] The account of equality presented in this Chapter affirms and helps ground these claims.

---

[308] See also on principles as direct norms for action: van der Burg, 'The Morality of Aspiration: A Neglected Dimension of Law and Morality', 177.

[309] Or in the case of this dissertation's concluding section: policy choices made in the regulation of algorithmic decision-making systems.

[310] To see equality as a principle that can be used to evaluate a decision-making process, is different from using equality as a principle that can be integrated by-design, i.e., as a technical instruction within the decision-making process itself. The use equality as an evaluative lens offers more flexibility. As an evaluative mechanism, equality becomes a diagnostic tool through which issues within the larger data-driven algorithmic environment can be identified: equality can be used to assess the technical details underlying a decision-making process (e.g. are the data considered within an algorithmic process reflective of voluntary choice), to contextual and procedural questions surrounding their deployment (e.g. was a decision transparent enough so people can understand how their actions will be taken into account by a decision-making process?). To incorporate equality by-design is a far more complicated exercise. Indeed, if purely seen as an element of design, equality must first be formalised. Technical approaches run the risk however of oversimplifying the demands of justice or the law. There is moreover a risk that equality would be transformed into a purely procedural principle, stripped, and decontextualized of the substantive layers that make it valuable. While technical measures to enhance or preserve equality are laudable, they must be complemented with a more holistic understanding of equality. Articulations of equality could function as a review-mechanism. Once a technical process no longer corresponds to our conceptualisation of equality, we can correct it if necessary. See for instance also: Bettina Berendt and Sören Preibusch, 'Better Decision Support through Exploratory Discrimination-Aware Data Mining: Foundations and Empirical Evidence', *Artificial Intelligence and Law* 22, no. 2 (1 June 2014): 175–209, https://doi.org/10.1007/s10506-013-9152-0; Bettina Berendt and Sören Preibusch, 'Toward Accountable Discrimination-Aware Data Mining: The Importance of Keeping the Human in the Loop—and Under the Looking Glass', *Big Data* 5, no. 2 (1 June 2017): 135–52, https://doi.org/10.1089/big.2016.0055.

[311] Viljoen, 'A Relational Theory of Data Governance', 654.

Algorithmic decisions rely upon differentiation. Hence, they tend to introduce inequality, rather than maintain (strict) equality. This is inherent to any decision-making process though: most societies must therefore tolerate to some extent the existence of certain inequalities.[312] As a consequence, and with the exception that people ought to be treated with equal concern and respect, if algorithms are to become the standard way of decision-making, algorithmic inequalities, whether they affect people's relationships or their distributive shares, must be tolerated to some degree as well. The important question is therefore not whether algorithmic inequalities can be avoided (they cannot!), but to determine to what extent algorithmic inequalities can be justified.

In this Chapter, I have shown how algorithms can interfere with the realisation of these interests. More specifically, every time an algorithmic decision interferes with one of the identified dimensions of equality, it has the potential to interfere with the desired structure of society. Still, specific conceptualisations of equality will concretise what fundamental interests ought to be equalised within society and how this equalisation process can be achieved. They can thus inform the development, deployment or regulation of algorithmically guided decisions by specifying precisely when, and under what conditions, algorithmically generated inequalities constitute a harm that must be addressed.[313] Yet, once we uncover a decision- or regulator's approach to equality, we can also evaluate whether this approach will be sufficiently robust to recognise and address the moral, socio-relational and distributive harms historically disadvantaged and non-representational groups and their group-members may experience within the algorithmic environment.

Having identified the various normative ideals typically ascribed to equality as a social value, I proposed to approach equality as a multidimensional concept that can be specified alongside three interrelated axes. This normative model of equality has been used as a practicable, analytical, and interpretative lens to identify and articulate a series of broad egalitarian harms algorithmic decision-making processes risk to impose onto decision-subjects. In the following Chapters, I demonstrate how this model can be operationalised further to evaluate algorithmic injustice, and the response formulated thereto by a given policy, law, code, or theory. In Part II, I

---

[312] See also: John Rawls, *Justice as Fairness: A Restatement*, ed. Kelly Erin (Cambridge, Massachusetts ; London, England: Belknap Press: Harvard University Press, 2001), 55.

[313] Equality may however be one value among many. An algorithmic inequality might be identified as a harm, but overruled by other interests, such as the freedom to conduct business. In other instances, the egalitarian harm gains priority and must be addressed.

identify and articulate the interpretation given to equality in European equality and non-discrimination law and investigate whether this legal conceptualisation is able to address the risk of algorithmic egalitarian harm. In Part III, I further deepen our understanding of the specific egalitarian harms decision-subjects may experience and the circumstances under which these harms can manifest, and the institutional safeguards decision-subjects need so they can challenge unjust inequalities. To that end, I transpose specific socio-relational and distributive notions of equality - chosen due to their correspondence with the multidimensional model - onto the algorithmic environment. In case the law's response appears inadequate, these political philosophical notions may inform and complement (future) regulatory governance initiatives.

PART II – Evaluating the Law: The Legal Conceptualisation of Equality and the Risk of Algorithmic Egalitarian Harm.

Introduction

As a social value, equality represents a society's commitment to equalise among its citizens the enjoyment certain fundamental interests. Having identified a core set of social ideals commonly associated with the notion of equality, I defined equality as a multidimensional concept that can be specified alongside a moral (equal concern and respect), socio-relational (equal social standing) and/or a distributive (equal access to certain justice-relevant goods) dimension. By positioning this model against algorithms' capacity to restructure society, we successfully operationalised equality to help us articulate and identify a set of egalitarian harms algorithms risk to impose. I argued that these risks should be taken seriously.

Having identified how algorithms threaten equality, we can assess a given law, policy, or other regulatory initiative in light of its ability to address said threat. In order to do so, we need to first uncover the specific interpretation given to equality by the framework subject to our evaluation. Once uncovered, we will know when people have a valid claim to equality and to what end. This response in turn can be mapped onto the algorithmic environment. **In Part II, I** operationalise the multidimensional concept of equality in order to perform this evaluative exercise: through it, I identify and assess the legal conceptualisation of equality as defined by European equality and non-discrimination law.

I first explain how the role and function of equality within European non-discrimination law is best uncovered alongside two axes of inquiry **(Chapter 3)**. I then examine the function of equality in the Council of Europe (CoE) legal order, focusing on the position of equality as a fundamental human right enshrined within the European Convention of Human Rights (ECHR) and subject to the judicial interpretation of the European Court of Human Rights (ECtHR) **(Chapter 4).**[314] Finally, I investigate the position equality holds as a fundamental principle of European Union (EU) law **(Chapter 5).** During this analysis, I focus on the general principle of equality as enshrined within

[314] European Convention on Human Rights (1950) (ECHR); Art. 14 ECHR; Protocol No. 12 to the Convention for the Protection of Human Rights and Fundamental Rights (2000). The ECHR is moreover subject to judicial interpretation by the European Court of Human Rights (ECtHR), whose judges interpret the Convention as "a living instrument, to be interpreted in light of present-day conditions." Inze v. Austria App no 8695/79 (ECtHR, 28 October 1987) para 41. See also on the role of judges in the ECtHR's equality and non-discrimination case-law. Barbara Havelková, 'Judicial Scepticism of Discrimination at the ECtHR', in *Foundations of Indirect Discrimination Law*, ed. Hugh Collins and Tarunabh Khaitan, 1st ed. (Oxford: Hart Publishing, 2018), 83–104, http://www.bloomsburycollections.com/book/foundations-of-indirect-discrimination-law/ch4-judicial-scepticism-of-discrimination-at-the-ecthr/.

the Charter of Fundamental Rights of the European Union (CFR), expressed via the EU Equality Directives, and interpreted through the case-law of the European Court of Justice (CJEU).[315]

Both European frameworks, each in their own way, define the conditions under which people can be differentiated and the inequalities they can justifiably face. As aforementioned, my objective is twofold. In a first step, I want to uncover the egalitarian objectives both orders pursue via their interpretation of equality. In a second step, and in parallel, I demonstrate how the normative choices made by the European legislator and judiciary in their interpretation of equality may either constrain or enhance the law's ability to address the egalitarian harms algorithmic decisions risk to impose.

[315] Consolidated Version of the Treaty on the Functioning of the European Union (2012) OJ C-326/47 (TFEU); Articles 8, 18 – 20 in particular; Consolidated Version of the Treaty on European Union (2012) OJ C-326/13 (TEU), Articles 2, 3 (3) and 9. Charter of Fundamental Rights of the European Union (2000) OJ C-361-1 (CFR); Chapter 3, Article 20 and 21 in particular. They are: Council Directive 2000/43/EC of 29 June 2000 implementing the principle of equal treatment between persons irrespective of racial or ethnic origin; Available via: https://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=CELEX:32000L0043:en:HTML; Council Directive 2000/78/EC of 27 November 2000 establishing a general framework for equal treatment in employment and occupation; Available via: https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=celex%3A32000L0078; Directive 2006/54/EC of the European Parliament and of the Council of 5 July 2006 on the implementation of the principle of equal opportunities and equal treatment of men and women in matters of employment and occupation (recast); Available via: https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A32006L0054; Council Directive 2004/113/EC of 13 December 2004 implementing the principle of equal treatment between men and women in the access to and supply of goods and services; Available via: https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=celex:32004L0113; Proposal for a Council Directive on implementing the principle of equal treatment between persons irrespective of religion or belief, disability, age or sexual orientation {SEC(2008) 2180} {SEC(2008) 2181}; available via: https://eur-lex.europa.eu/legal-content/en/TXT/?uri=CELEX%3A52008PC0426.

3 Uncovering The Legal Conceptualisation of Equality

Assisted by my multidimensional model of equality, and in light of the risks already identified in Part I, in this Chapter, I argue that the role and function of equality within European non-discrimination law is best uncovered alongside two axes of inquiry. In following two Chapters, these axes of inquiry will focus my navigation of the legal landscape.

Before I turn to these axes of inquiry, I first introduce two overarching contextual factors that may affect the interpretation of equality as a legal concept: a) the idea that the notion of equality can hold various different functions within the same legal order and b) how in these functions, equality can be treated either as a procedural and formal notion or as a goal-oriented and substantive notion. **(Section 3.1).** I then discuss the first area of inquiry, which investigates whether the law follows an open- or closed-model approach to equality **(Section 3.2.1).** An exploration of this research question allows to uncover *whom* the law protects against unjust inequality. The second area of inquiry concerns the capacity for decision-makers to justify differential treatment **(Section 3.2.2).** The latter inquiry can be further specified. In order to uncover the function equality holds within a legal order, insight must first be gained into the review mechanisms applied when equality is positioned against competing values. Second, when this happens, what particular ideals does equality represent? Do decision-subjects have the ability to have their moral, socio-relational, or distributive egalitarian interests recognised? Through an exploration of these axes of inquiry, it can be ascertained which algorithmic inequalities a particular (legal) conceptualisation of equality can, or is willing, to address. Finally, I compare the (predominantly) grounds-based nature of both European legal systems against an alternative, more principle-oriented approach to equality **(Section 3.2.3).**

3.1 Equality: A Multi-Layered Legal Notion

In the following paragraphs, I want to highlight two types of distinctions that can help shape our understanding of equality as a legal principle, and the various functions equality can play as a legal principle. First, a distinction can be made between various models of equality as a legal principle: the second distinction concerns the evolution of equality from a formal notion, to one that is more substantive in nature.

3.1.1 Models of Equality

In the previous Part, equality was defined as a concept that represents three different ideals. First, a given policy or framework affirms equality as a moral ideal when it takes into account the interests of people, so that every person's interests matter and matter equally, Second, as a socio-relational ideal, the notion of equality is linked to the social relationships people stand in. An investigation into the socio-relational dimension of a given policy clarifies the efforts that are made to ensure that people can maintain relationships of equal standing. Finally, by investigating a policy's distributive dimension, we can identify the(non-)material justice-relevant goods people should have (roughly) equal access to.

Legal scholars have observed how equality can also hold several functions within the legal order. These functions can incorporate elements of, but do not always correspond to the aforementioned abstract model.[316] Sottiaux for example identifies three core models of equality as a legal concept.[317] Under a first model, equality acts as a principle of good governance.[318] Here, reference to equality serves to scrutinise actions where differential treatment is imposed onto people, situations or cases, but for which no reasonable justification can be provided. Hence, equality is seen as an expression of the requirement to be either reasonable or rational. Such an approach appears neutral. It simply serves to investigate whether a measure can be justified or not: if a decision-maker can provide no concrete justification as to why a measure was necessary, the action was unreasonable or arbitrary and therefore unjustified. This model of equality can be applied to every type of differential treatment - albeit that it is traditionally understood as a vertical principle that targets governmental and public sector bodies.[319] Hence, equality acts as an institutional benchmark. Under the second model, equality features as a human right.[320] Here equality attains a clear normative function: to combat and ban certain forms of differential treatment experienced as denigrating, exclusionary, or otherwise harmful. This human rights function is typically given shape by prohibiting differentiation on the basis of (exhaustively enumerated) suspect characteristics or grounds. When this function is ascribed to equality, the

---

316 Stefan Sottiaux, 'Het Gelijkheidsbeginsel : Langs Oude Paden En Nieuwe Wegen (Artikel, 2008) [WorldCat.Org]', *Rechtskundig Weekblad* 72, no. 17 (27 December 2008): 690–702.

317 Sottiaux, 690. Spinoy observes however that See also: Marie Spinoy, 'Discriminatie in Het Gelijkekansendecreet? Reflecties Bij GwH Nr. 110/2019', *Tijdschrift Voor Bestuurswetenschappen En Publiekrecht*, no. 7 (2020): 337.

318 In Dutch: "beginsel van behoorlijke politieke besluitvorming." Sottiaux, 'Het Gelijkheidsbeginsel : Langs Oude Paden En Nieuwe Wegen (Artikel, 2008) [WorldCat.Org]', 691.

319 Sottiaux, 691.

320 Sottiaux, 695. See also : Spinoy, 'Discriminatie in Het Gelijkekansendecreet? Reflecties Bij GwH Nr. 110/2019'.

notion tends to operate on both a vertical and horizontal level but can often only be invoked in relation to specific prized goods. Finally, and rather than an instrument for people or groups to combat certain forms of disadvantageous treatment, equality can be seen as a proactive policy goal that encourages governments or decision-makers to realise social equality, inclusivity, and diversity and ensure fair or *real* equality of opportunity.[321]

Traces of every model can be found in most legal orders.[322] The identification of the model at play can help calibrate people's expectations concerning the law's response to algorithmic inequalities. Following the distinction above, where equality functions as a human right or a proactive tool towards diversity, it is typically infused with a concrete socio-relational or distributional purpose in mind. Equality as a principle of good governance is not explicitly value laden. A given measure

[321] Sottiaux, 'Het Gelijkheidsbeginsel : Langs Oude Paden En Nieuwe Wegen (Artikel, 2008) [WorldCat.Org]', 699. In a slightly different model of equality, McCrudden and Kountouros identify four approaches towards equality as a fundamental right. Under equality as rationality, equality is a "self-standing principle of general application, without specific limitation on the circumstances in which it is applicable (except that it be in the public realm, broadly defined), and without limitation on the grounds on which the difference of treatment is challengeable." The second approach concerns equality as protective of prized goods, whereby certain goods are identified as being so important that their distribution should be guaranteed to all. Under a third approach, equality can be seen as preventing "status-harms" that can flow from differential treatment on particular protected grounds. Finally, equality may be approaches as constituting a proactive duty to promote equal opportunity between particular groups. Christopher McCrudden and Haris Kountouros, 'Human Rights and European Equality Law', in *Equality Law in an Enlarged European Union: Understanding the Article 13 Directives*, ed. Helen Meenan (Cambridge: Cambridge University Press, 2007), 74–78, https://doi.org/10.1017/CBO9780511493898.004. See also: Christopher McCrudden and Sacha Prechal, 'The Concepts of Equality and Non-Discrimination in Europe: A Practical Approach' (European Network of Legal Experts in the Field of Gender Equality, November 2009).

[322] Furthermore, depending on the model at play, the actors who should uphold the duty to guarantee equality may differ, i.e., should equality be guaranteed by public authorities, private entities, or both? Second, the level of scrutiny applied by the Courts vis-à-vis differential treatment may differ depending on the model of equality at play as well. While a government has to balance between a variety of interests of citizens and is typically given a high degree of freedom in doing so (equality as good governance/low level of scrutiny), they will face greater difficulty in justifying a policy where it is based upon a suspect ground (equality as a human right/higher level of scrutiny). At the same time, models of equality are not isolated from one another, nor are they static: they interact and evolve, whereby the different models can inform or contradict one another. In this regard the evolution of equality as a human right can inform the application of equality as a principle of good governance. Spinoy, 'Discriminatie in Het Gelijkekansendecreet? Reflecties Bij GwH Nr. 110/2019'. Sottiaux, 'Het Gelijkheidsbeginsel : Langs Oude Paden En Nieuwe Wegen (Artikel, 2008) [WorldCat.Org]', 690; Spinoy, 'Discriminatie in Het Gelijkekansendecreet? Reflecties Bij GwH Nr. 110/2019', 337.

might still be evaluated in light of the socio-relational or distributional harms it imposes, but those harms are weighed against the interests of, and justification provided by, the decision-makers.[323]

3.1.2 Formal and Substantive Approaches Towards Equality

Beyond the models of equality, a second distinction that is often made with regard to equality law, is the distinction between formal and substantive equality.[324]

FORMAL APPROACHES – Formal approaches towards equality depart from the premise that like situations should be treated alike.[325] Under a formal view, a comparison will take place: if two situations are comparable or analogous, they ought to be treated alike, and if they are not comparable, they ought to be treated differently.[326] Formal equality has been described as a, symmetrical and (open) anti-classification approach towards equality grounded in ideals of individual justice and merit.[327] Under a formal view, any classification or act of differentiation could be met with suspicion.[328] Whenever a measure is thought to deviate from the basic premise that like situations should be treated alike, it can be made subject to scrutiny. As such, any form of

---

[323] Sottiaux refers to this treatment as constituting no moral judgment in the sense that one does not need to illustrate an intent to harm through differential treatment. However, the justification process involved is not without consequence, and could also have normative implications. Tridimas rightfully observes how the definition of arbitrary conduct will be built upon ideas that are not politically neutral. For instance, in the case of EU-law, distributive justice and classic liberalism might function as guiding ideas. Tridimas, *The General Principles of EU Law*, 62.

[324] In relation to the aforementioned distinction, Sottiaux has observed how a formal notion of equality underlies the models of equality as good governance and equality as a human rights instrument. Positive or affirmative equality measures, as well as measures aimed towards the protection of diversity are more substantive in nature. The distinction has also been described as *de jure* versus *de facto* equality. See Olivier De Schutter, 'Three Models of Equality and European Anti-Discrimination Law', *Northern Ireland Legal Quarterly* 57, no. 1 (2006): 1. Sottiaux, 'Het Gelijkheidsbeginsel : Langs Oude Paden En Nieuwe Wegen (Artikel, 2008) [WorldCat.Org]', 690.

[325] See also Croon-Gestefeld for a more nuanced version regarding the relationship between formal and Aristotelian equality. Croon-Gestefeld, *Reconceptualising European Equality Law: A Comparative Institutional Analysis*, 30–31. However, to some the idea that "differences should be treated differently" is already one step towards a more richer, and substantive understanding of equality. McCrudden and Prechal, 'The Concepts of Equality and Non-Discrimination in Europe: A Practical Approach'.

[326] Marc De Vos, 'Substantive Formal Equality in EU Non-Discrimination Law', in *The European Union as Protector and Promoter of Equality*, ed. Thomas Giegerich, European Union and Its Neighbours in a Globalized World (Cham: Springer International Publishing, 2020), 246, https://doi.org/10.1007/978-3-030-43764-0_13.

[327] De Vos, 246. The promotion of individual autonomy can also be seen as part of these liberal values. Sandra Fredman, 'Equality: A New Generation', *Industrial Law Journal* 30 (2001): 154.

[328] Sandra Fredman, 'Providing Equality: Substantive Equality and the Positive Duty to Provide', *South African Journal on Human Rights* 21 (2005): 163.

differential treatment could trigger an investigation. The approach is procedural in the sense that it triggers an investigation into the consistency of a decision-makers' treatment of its decision-subjects.[329]

Because differentiation is necessary in practice however, when following a formal approach, the law is likely to define certain characteristics or traits as "irrelevant" or "suspect" for decision-making, such as a person's sex or ethnicity. In principle, differential treatment cannot be justified when these traits are used. In these cases, the law offers symmetrical protection: all classifications based upon those traits are suspicious and everyone is equally protected against their use.[330] As the use of these traits is to be avoided, under a formal view, abstraction is made of certain identity markers; they are rendered invisible. In so doing, people's individual actions, behaviour, or merit are prioritised.[331] Yet, this agnostic approach towards identity has been criticised because it tends to ignore the wider social and economic context people find themselves in, such as the institutional structures they are governed by or the groups they are a member of.[332]

SUBSTANTIVE APPROACHES – Formal approaches towards equality are often contrasted with substantive ones: "[Substantive equality] moves beyond or away from individual merit-based comparison to account for the broader context that affects personal merit, it is dynamic in its ambition to influence that context and improve desired outcomes, internalizing various collective or societal dimensions along the way."[333] When equality is viewed as a substantive notion, we are not necessarily concerned with the identification of cases that are similar or different. Instead, equality becomes a notion with a specific socio-relational and distributional purpose. For instance, Fredman's influential account of substantive equality identifies four key goals equality laws should address: "to redress disadvantage; to address stigma, stereotyping, prejudice and violence; to enhance voice and participation; and to accommodate difference and achieve structural change."[334] Whereas formal frameworks define certain traits as irrelevant, under a substantive approach these traits gain meaning.[335] For example, Fredman observes how a

---

[329] Sandra Fredman FBA, *Discrimination Law*, 2nd ed. (Oxford ; New York: Oxford University Press, 2011), 9.

[330] Fredman, 'Providing Equality', 164.

[331] See also: Fredman, 'Substantive Equality Revisited', 732–33.

[332] Marc De Vos, 'The European Court of Justice and the March towards Substantive Equality in European Union Anti-Discrimination Law', *International Journal of Discrimination and the Law* 20, no. 1 (1 March 2020): 75, https://doi.org/10.1177/1358229120927947.

[333] De Vos, 63–64.

[334] Fredman, 'Substantive Equality Revisited'.

[335] Fredman, 733.

person's gender or ethnicity is not necessarily irrelevant; they can constitute valued components of a person's identity or sense of self: "the problem is not so much difference per se, but the detriment which is attached to difference." Even though people of colour have been treated in discriminatory ways due to their ethnicity, ethnicity can still be an important identity-marker. This difference must be recognised, rather than rendered abstract: why prohibit the use of one's ethnicity as a differentiating factor if its use can improve the position of those typically disadvantaged due to their ethnicity?[336]

EU EQUALITY LAWS - Historically, both ECHR and EU equality and non-discrimination frameworks share a formal foundation towards equality.[337] Today however, they represent a hybrid approach, whereby elements of both formal and substantive equality can be discerned.[338] Both approaches can operate in parallel to one another and both have their advantages. For instance, if sufficiently made explicit, a substantive conceptualisation of equality will allow a clearer delineation of those (algorithmic) inequalities that are unjustified. At the same time, a formal approach allows for some flexibility.[339] Because any instance of differentiation can be treated as suspicious, it allows for an investigation of any algorithmic decision-making act.[340] The latter can be advantageous in

---

[336] Fredman, 'Providing Equality', 164.

[337] Christa Tobler, 'Equality and Non-Discrimination under the ECHR and EU Law', 2014, 529.Muir notes how expressions of equality under EU law find a common origin in the Aristotelian formula. Elise Muir, 'The Essence of the Fundamental Right to Equal Treatment: Back to the Origins', *German Law Journal* 20, no. 6 (September 2019): 817–18, https://doi.org/10.1017/glj.2019.64. De Vos, 'The European Court of Justice and the March towards Substantive Equality in European Union Anti-Discrimination Law'. See for instance: Sandra Fredman, 'Emerging from the Shadows: Substantive Equality and Article 14 of the European Convention on Human Rights', *Human Rights Law Review* 16, no. 2 (16 April 2016): 279, https://doi.org/10.1093/hrlr/ngw001.

[338] For example, the distinction between direct and indirect discrimination, used in both European systems, is often seen as being representative of the formal and substantive dichotomy as direct discrimination focuses on disadvantageous treatment (on the basis of a prohibited characteristic) and indirect discrimination focuses on the disadvantageous impact a given measure might have for people belonging to a protected group. The notion of indirect discrimination is therefore more likely capture institutional forms of disadvantage. Still, it is not entirely substantive as the obligation to avoid indirect discrimination does not impose a duty upon decision-makers to actively change structures in order to ensure that disadvantage is institutionally rectified. See for instance: De Vos, 'The European Court of Justice and the March towards Substantive Equality in European Union Anti-Discrimination Law'; De Vos, 'Substantive Formal Equality in EU Non-Discrimination Law'; Fredman, 'Substantive Equality Revisited'; Catherine Barnard and Bob Hepple, 'Substantive Equality', *Cambridge Law Journal*, no. 59 (21 December 2000): 562–85.

[339] Albeit that entirely formal approaches will continue to showcase deficits considering algorithmically guided decision-making processes. For instance, the comparability requirement is one area where formal approaches will run into difficulties. See Chapter 4, Section 4.3 "Comparability"

[340] Furthermore, even under a formal approach to equality, one could take into consideration more substantive ideals as part of the review process. For example, the reasons a decision-maker must advance to justify a given act could

an environment hallmarked by technologies that enable differentiation between a great variety of groups on a whole range of grounds for a diverse set of purposes.[341] Yet despite their purported neutrality, formal approaches towards equality are far from impartial. In ruling whether two situations are comparable or different, and whether differential or similar treatment can be objectively justified, certain choices have to be made by the assessor and these choices can affect the outcome of a given case, *in casu* whether a given harm can find protection or not.[342]

3.2 The Legal Conceptualisation of Equality: Two Lines of Inquiry

In the following sub-Sections I argue that the legal conceptualisation of equality can be uncovered through two broad lines of inquiry. The first line of inquiry investigates whether a system maintains an open or closed approach towards equality and non-discrimination claims. The second line of inquiry concerns the possibility for decision-makers to justify differential treatment. This investigation can be linked to the three-dimension model of equality, as well as our observations concerning the question "Equality of Whom?", presented in the previous Chapter. More specifically, this investigation will allow us to identify what dimensions of equality both systems look to protect and who can benefit from this protection. Those answers however

become stricter where that act disproportionately targets a historically disadvantaged group. See also: De Vos, 'The European Court of Justice and the March towards Substantive Equality in European Union Anti-Discrimination Law'; De Vos, 'Substantive Formal Equality in EU Non-Discrimination Law'.

[341] An alternative distinction towards equality is presented by Gerards. Rather than a formal or substantive approach towards equality, the law can be defined by the choice between using the principle of non-discrimination as a foundation, or the principle of equality or equal treatment. Under a non-discrimination rationale, equality serves to protect individuals and groups against differential treatment upon the basis of certain suspect grounds, and the irrelevance or irrationality that their use represents. Under the equal treatment rationale, equality has an instrumental or procedural function, whereby it allows a person "[…] to demonstrate arbitrariness or unfair treatment by pointing at other cases, similar to [their] own case, where more favourable decisions are taken, or which are governed by a less burdensome legal regime." As a consequence, each difference would need to be assessed in terms of its reasonableness or fairness by a reviewing party, i.e., a Court: "The only relevant question is if one group or person is allowed to exercise a certain right or receive a certain benefit, whilst this is not permitted for another person or group." In this context, the equal treatment rationale would serve as a flexible complement for the evaluation of algorithmically guided decision-making processes. J. Gerards, 'The Discrimination Grounds of Article 14 of the European Convention on Human Rights', *Human Rights Law Review* 13, no. 1 (1 March 2013): 113–17, https://doi.org/10.1093/hrlr/ngs044.

[342] See also: T.W.A. de Wit, 'Wanneer Iedereen Aan Zijn Eigen Waarheid de Voorkeur Geeft. Kan Politiek Wel Zonder Common Ground?', in *Onenigheid En Gemeenschap. Basisboek Politieke Filosofie*, ed. M. Terpstra (Amsterdam: Boom, 2012), 528–48. Formal approaches to equality are therefore not normatively empty, as some authors have claimed. Westen, 'The Empty Idea of Equality'.

might drastically change depending on what model of equality, as well as approach towards equality, formal or substantive, is at play. [343]

3.2.1 The (Initial) Range of Protection: Open or Closed Models of Equality?

Should any instance of differentiation give rise to an equality claim? This question lies at the heart of a legal system's choice between an open or closed approach towards equality and non-discrimination. Because both European legal systems tend to define equality as the absence of discrimination, the question can be rephrased as follows: can any form of differentiation amount to (a prohibited form of) discrimination or did the law place *definitional* constraints onto the notions of inequality and discrimination *a priori*.[344]

OPEN MODELS - Open models do not define beforehand what forms of differential treatment will constitute a discriminatory act or violation of equality.[345] The open model towards equality is procedural and instrumental in the sense that any differentiating act, and the inequalities that follow thereout, could constitute a harm and should therefore be made subject to evaluation.[346] Any person (or group) who believes they have been unfairly differentiated against can invoke the principle of equality or non-discrimination. To the extent that the open model has the potential to treat any form of differentiation as suspicious, it represents a formal approach. Whether a differentiating amounts to unjustifiable discrimination, is a question that will be resolved during the review or justification phase.

CLOSED MODELS - Closed models delimit the protective scope of the law *a priori.* They put forward a rule that will define what acts or rules can constitute discrimination.[347] For instance, scrutiny

---

[343] For an alternative evaluation of European non-discrimination models in light of algorithmic decision-making, see in particular: Gerards and Zuiderveen Borgesius, 'Protected Grounds and the System of Non-Discrimination Law in the Context of Algorithmic Decision-Making and Artificial Intelligence', June 2022.

[344] See also: Rory O'Connell, 'Cinderella Comes to the Ball: Art 14 and the Right to Non-Discrimination in the ECHR', *Legal Studies* 29 (2009): 211–29.

[345] Oddný Mjöll Arnardóttir, *Equality and Non-Discrimination under the European Convention on Human Rights* (Brill Nijhoff, 2021), 33, http://brill.com/view/title/8988.

[346] Gerards, 'The Discrimination Grounds of Article 14 of the European Convention on Human Rights'.

[347] Arnardóttir, *Equality and Non-Discrimination under the European Convention on Human Rights*, 33. In reference to: Aalt Willem Heringa, 'Standards of Review for Discrimination. The Scope of Review by the Courts.', in *Non-*

can only be triggered where a differentiating act can be tied or linked to a suspect or prohibited characteristic. Legal protection is made conditional and available only to those acts or practices that make it through an initial filtering mechanism. In so doing, closed models restrict the goals, functions and limits of equality and non-discrimination can achieve.[348]

CHOOSING BETWEEN MODELS – There are several reasons why one would choose a closed model over an open one. First, closed models can represent and achieve a (substantive) vision on equality more directly. For instance, one could argue that the law should prioritise the protection of historically disadvantaged groups against discrimination and inequality. In so doing, the closed nature of a legal system might also gain an expressive and symbolic function as it clearly delineates those acts of discrimination that are socially reprehensible. Under an open approach, time and resources would be allocated to scrutinise inequalities that are less socially pressing. Second, the choice for a closed model might be prompted by practical and legal considerations, such as the fear for case-law inflation or legal uncertainty. At the same time, an open model might have the benefit that it can more easily treat equality as an evolving concept and can accommodate changing conceptions as to what types of inequality and differentiation are particularly nefarious or unjust.[349]

CHOICE OF MODELS AND ALGORITHMS - From an algorithmic perspective open models have one clear benefit. Algorithmic differentiation acts and associated inequalities come in many different forms and guises. The algorithmic environment presents us with a great degree of uncertainty regarding what forms of differential treatment are justified and those that are not. Under an open model, all these discrete instances of differentiation could be analysed and investigated on a case-by-case basis.[350] Because every algorithmic differentiation ground tends to capture an aggregate of people,

---

Discrimination Law: Comparative Perspectives, ed. Titia Loenen and Peter R. Rodrigues (The Hague: Kluwer Law International, 1999), 27.

[348] The distinction between models need not be as sharply drawn. For example, whereas law makers might have chosen for an open-ended equality clause, over time, and even though every person could continue to advance their equality claims, the open clause's protective scope might have been narrowed down due to judicial interpretation.

[349] Naudts, 'Criminal Profiling and Non-Discrimination: On Firm Grounds for the Digital Era?'; Naudts, 'AI En Algoritmische Categorisatie: Gelijkheid En Non-Discriminatie'. Gerards and Zuiderveen Borgesius, 'Protected Grounds and the System of Non-Discrimination Law in the Context of Algorithmic Decision-Making and Artificial Intelligence', 2 November 2020, 38–41.

[350] For arguments in favour of an open approach: Naudts, 'AI En Algoritmische Categorisatie: Gelijkheid En Non-Discriminatie'; Gerards and Zuiderveen Borgesius, 'Protected Grounds and the System of Non-Discrimination Law in the Context of Algorithmic Decision-Making and Artificial Intelligence', 2 November 2020.

any person, regardless of their current social position or their possession of a given characteristic, could ask for a review of the egalitarian harm they have incurred.

While this sounds promising, the algorithmic environment also presents arguments against such an open-ended approach. Algorithmic processes have transformed the way in which decisions are made. They have moved decision-making away from causal-based to correlation-based reasoning. Under an open model, decisions on how to deal with these evolutions would be placed squarely in the hands of judicial authorities. Yet, questions could be raised regarding the legitimacy of such judicial decision-making power as it moves fundamental questions concerning the structuring of society away from public discourse.[351]

3.2.2 The Justification of Differentiation

Regardless of whether the law has an open or closed approach towards equality, not every instance of differentiation can be prohibited. Differentiation is a necessary component of decision-making, and any given society must tolerate the existence of certain inequalities.

Both European legal systems have therefore developed certain methods or mechanisms in order to delineate between those acts of differentiation that are allowed and those that are not. These mechanisms can be directly provided by the legislature. The EU Equality Directives for instance stipulate when, and under what conditions, an act of differentiation constitutes a prohibited form of discrimination and hence an unjustified inequality. The justification of differentiation is often further concretised through the case-law of the ECtHR and CJEU, either in their interpretation of substantive non-discrimination laws, or in their interpretation of equality as a fundamental legal principle or human right.

By looking more concretely at how a legal system's justification procedure is structured, one can uncover what interests a legal system is willing to protect, when and under what conditions. This investigation should be further specified in order to successfully assess the law's ability to address

[351] Eder rightfully made this observation. He put it as follows: "The right to equal treatment [an open approach] does not define the logic, on the basis of which comparisons and differentiations should be assessed but presupposes it." Niklas Eder, 'Privacy, Non-Discrimination and Equal Treatment: Developing a Fundamental Rights Response to Behavioural Profiling', in *Algorithmic Governance and Governance of Algorithms: Legal and Ethical Challenges*, ed. Martin Ebers and Marta Cantero Gamito, Data Science, Machine Intelligence, and Law (Cham: Springer International Publishing, 2021), 44, https://doi.org/10.1007/978-3-030-50559-2_2.

the egalitarian harms algorithms risk to impose. First, there is a need to investigate the structure of the review process. When unjust inequalities are challenged, how will this evaluation typically take place: Is the equality review multi-tiered, and if so, what is the evaluative focal point of each tier? Are there different justification procedures, and if so, when do they apply and to what extent do they differ? Second, whether equality is governed through a legislative instrument or the judiciary, it represents a social value that competes with others, such as individual liberty or the freedom to conduct a business. Still, such a balancing exercise can only be accurately performed where the interests in question have been articulated. One would therefore expect that the justification procedure involves (and also allows) the identification of those egalitarian interests a given legal order wants to secure, and those egalitarian disadvantages it wants to guard people against. The (normative) role and function equality holds within a specific legal order should be revealed as part of the justification process. The dimensional lens can help to further navigate, identify, and articulate the egalitarian ideals a given legal system (explicitly or implicitly) protects. For instance, where it deems a discriminatory act unjust, is this conclusion reached on the basis of socio-relational or distributive paradigm?

3.2.3 Grounds or Principles?

Beyond the two lines of inquiry central to the conceptualisation of equality, a clarification can be made about the subject or focus of evaluation in cases of inequality and discrimination cases.

GROUNDS-BASED APPROACH - Both European systems are (predominantly) grounds-based. Grounds-based equality models emphasise the characteristics that are used as a basis for differentiation. Typically, lawmakers will define a set of suspect characteristics, the use of which is either prohibited or met with a heightened degree of suspicion. The reason certain grounds are met with heightened caution can be varied. Formal approaches treat grounds as suspect when they are irrelevant. Under a substantive account, the suspect grounds tend to be more symbolic: they are seen as a proxy for the disadvantage, prejudice or stigma associated with the possession of a given trait.[352] In any case, during the evaluation of equality claims the ground used to

[352] See for example: Khaitan, *A Theory of Discrimination Law*; Elisabeth Holzleithner, 'Mainstreaming Equality: Dis/Entangling Grounds of Discrimination Symposium: Antidiscrimination Law in the European Union', *Transnational Law & Contemporary Problems* 14, no. 3 (2005 2004): 934; Nicholas Bamforth, 'Sexual Orientation Discrimination after Grant v South-West Trains', *Modern Law Review* 63, no. 5 (September 2000): 702, https://doi.org/10.1111/1468-2230.00287. A system that is built around the protection of innate personal traits may reflect a society's dignitarian commitment to protect people from being adversely treated on the basis of characteristics that relate to the core of their identity and personality, such as their religion, political opinion or sexuality. Bamforth, 702.

differentiate between people tend to play a decisive role.[353] Grounds gain even greater importance under a fully closed system: unless a discriminatory act or inequality can be (directly or indirectly) tied to a specific characteristic, no protection can be obtained.[354]

Under a grounds-based model, one should also decide whether discrimination claims are demarcated by a single ground (single-axis models), or whether discrimination claims are open to more complex forms of differentiation. The latter concern instances where disadvantage is experienced alongside multiple dimensions or at the intersections thereof. From both a societal and algorithmic perspective, one can caution against the (sole) use of simplistic, single-axis models.[355] Indeed, intersectional and complex forms of (algorithmic) differentiation can be more easily captured where a grounds-based approach recognises the multidimensional nature of certain inequalities.[356]

If the algorithmic debate, and the question whether differentiation was fair or unfair in particular, would be settled by identifying those differences that make a difference, a grounds-based approach would do exactly that. It would specify those circumstances under which an algorithmically guided decision cannot make use of a given ground, trait, or characteristic. In order to be workable such an approach must be sufficiently sound. Malleson for example contends that

---

[353] See for example: Gerards, 'The Discrimination Grounds of Article 14 of the European Convention on Human Rights', 119–21. Janneke H. Gerards, 'Intensity of Judicial Review in Equal Treatment Cases', *Netherlands International Law Review* 51, no. 2 (September 2004): 135–83; Oddný Mjöll Arnardóttir, 'The Differences That Make a Difference: Recent Developments on the Discrimination Grounds and the Margin of Appreciation under Article 14 of the European Convention on Human Rights', *Human Rights Law Review* 14, no. 4 (1 December 2014): 647–70, https://doi.org/10.1093/hrlr/ngu025; Oddný Mjöll Arnardóttir, 'Vulnerability under Article 14 of the European Convention on Human Rights', *Oslo Law Review* 4, no. 03 (15 December 2017): 150–71, https://doi.org/10.18261/issn.2387-3299-2017-03-03; Arnardóttir, *Equality and Non-Discrimination under the European Convention on Human Rights*; Nicholas Bamforth, 'Prohibited Grounds of Discrimination under EU Law and the European Convention on Human Rights: Problems of Contrast and Overlap', *Cambridge Yearbook of European Legal Studies* 9 (ed 2007): 1–42, https://doi.org/10.5235/152888712802746920; Aileen McColgan, 'Defining the Protected Characteristics', in *Discrimination, Equality and the Law*, 1st ed. (London: Hart Publishing, 2014), 38–69, http://www.bloomsburycollections.com/book///; Croon-Gestefeld, *Reconceptualising European Equality Law: A Comparative Institutional Analysis*.

[354] The relevance of the ground may differ depending upon what model of equality is at play, however. Where equality serves an institutional benchmark, there is often no *a priori* constraint on what differentiation ground can be made subject to scrutiny. Within the same legal system there might however be a specific non-discrimination law that does impose constraints on what grounds can be made subject to protection beforehand.

[355] See also: Shreya Atrey, *Intersectional Discrimination, Shreya Atrey.* (Oxford, England: Oxford University Press, 2019).

[356] See also Chapter 2, Section 2.3 "Equality and Intersectional Differentiation."

a framework based upon the identification of suspect grounds should satisfy three conditions. First, the framework must have definitional and categorial stability. Second, the system must broadly reflect people's understanding of social reality and their lived experiences. Third, the chosen grounds must correspond to the most significant axes of discrimination and disadvantage within society.[357] These three conditions are difficult to satisfy in an environment where social configurations of identity are under question or subject to change.[358] Whereas such shifts take place outside the digital realm, algorithms too function in ways that challenge the aforementioned conditions: they reconfigure the relationships people stand in and the identities attached to people on a continuous basis.[359] While people might experience disadvantage following an algorithmic decision, the impact thereof is not always immediately apparent from the differentiation ground in question.

PRINCIPLED APPROACH - Alternative evaluation methods that are more contextual or principled-driven could be explored to mitigate or complement the deficiencies of a purely grounds-oriented approach. As aforementioned in the previous Chapter, such an approach could instead focus on the disadvantage (e.g., socio-relational, or distributional harms) algorithms might impose onto individuals and groups of people. Conceptualisations of equality could help to define in more abstract terms what inequalities algorithmic decision-makers should not impose onto people and groups due the potential disadvantage they might introduce. For instance, departing from a socio-relational conceptualisation of equality, a principled approach could stipulate that algorithmic differentiation acts should not lead to the social exclusion of those most disadvantaged in society. This system would be closed, but the designation of persons who can launch an inequality claim would no longer tied to a specific trait or ground. This rule can adapt to changes in society and ensure that those most disadvantaged always benefit from protection.

---

[357] Kate Malleson, 'Equality Law and the Protected Characteristics', *The Modern Law Review* 81, no. 4 (2018): 598–621, https://doi.org/10.1111/1468-2230.12353.

[358] For example, in relation to the stability requirement, Malleson notes the following: "A grounds-based anti-discrimination law requires stable categories to function. If we continue to move into a post-racial, post-gender, post-binary world this may become increasingly problematic." Malleson, 609.

[359] Within society at large, these three conditions - and as a consequence grounds-based systems - are under pressure due to three trends: "the emergence of a much wider range of identity characteristics, the growing importance of intersectionality (both conceptually and experientially) and the increasing tendency to reconstruct identity characteristics along a spectrum." Kate Malleson, 'Equality Law and the Protected Characteristics: Equality Law and the Protected Characteristics', *The Modern Law Review* 81, no. 4 (July 2018): 598–621, https://doi.org/10.1111/1468-2230.12353.

A plurality of foundational principles can substantiate grounds-based systems. Still, when the chosen approach is too keenly centred on the grounds as such, it runs the risk of decontextualising the evaluation of differential treatment and associated inequalities from the harms society wanted to avoid in the first place..[360] If our main concern is to protect people against a specific disadvantage or harm, why tie people's protection against said harm to the presence of a specific differentiation ground? Moreover, rather than looking at a concrete harm or disadvantage a person or group might experience, the assessment of discriminatory treatment might be reduced to a formal exercise of spot-the-protected-characteristic.[361]

3.3 Conclusion

This Chapter laid the foundation for the legal analysis that will be performed in the next Chapters. More specifically, I have argued that a legal system's conceptualisation of equality as an ideal, and response to algorithmic inequality can be uncovered through two specific lines of inquiry. First, what inequalities a legal system is willing to address, will be dictated by a system's choice between an open or closed model towards differentiation. Unlike open models, closed ones place *a priori* constraints on what differentiation acts - and within grounds-based systems, differentiation grounds - can constitute discrimination and violate equality. Because algorithmic differentiation grounds capture people and groups of people, this line of inquiry bears significance on the "equality of whom?" debate. Open models allow anyone to launch a discrimination/equality claim when they believe they have been disadvantaged, whereas closed models impose a filtering mechanism. Second, and whatever the choice of system may be, the justification procedure will further specify the role and function equality holds within a legal order. While the justification procedure places equality against competing interests, it should highlight those egalitarian interests a given society wishes to defend and it should clarify under what circumstances those interests can be overridden. Ideally then, this exercise brings to the surface those dimensions of equality the law prioritises and how they can find protection.

---

[360] See also Chapter 2: Section 2.2.2 Algorithmic Equality of Whom?

[361] That is not to say that differentiation ground has no significance at all. Grounds could still maintain their signalling function. They could designate those inequalities that are particularly nefarious or morally reprehensible. See for example: Wibren van der Burg, 'The Expressive and Communicative Functions of Law, Especially with Regard to Moral Issues', *Law and Philosophy* 20, no. 1 (1 January 2001): 31–59, https://doi.org/10.1023/A:1006492921790; Sarah Ganty and Jua Carlos Benito Sanchez, 'Expanding the List of Protected Grounds within Anti-Discrimination Law in the EU: AN Equinet Report' (Equinet, 2021).

In the following two Chapters, and making use of the aforementioned lines of inquiry, I uncover the legal conceptualisation of equality found within European equality and non-discrimination laws by examining whom the law protects (equality of whom?) against which inequalities (what egalitarian ideals do they promote?) and how (what mechanisms have been put when equality is positioned against competing values?). In so doing, it will be possible to evaluate whether the law can address the broad range of egalitarian harms algorithms risk to impose onto a vast variety of people? [362]

[362] The end-goal is not to advance a foolproof method that can unequivocally predict how European equality and non-discrimination laws will respond to algorithmically guided decision-making practices and the inequalities they impose. Instead, by uncovering a legal system's approach towards equality, strong indicators as to their *likely* response will be given.

4 Equality and Non-Discrimination under the European Convention on Human Rights

Following the two critical areas of inquiry identified within Chapter 3, in Chapter 4, I examine the legal conceptualisation of equality found within the European Convention on Human Rights (ECHR) as given shape through the case-law of the European Court of Human Rights (ECtHR). After having provided a brief overview of the key equality and non-discrimination provisions (Section 4.1), the Chapter's second section will assess whether or not the ECtHR follows an open or closed approach towards their definition of equality and non-discrimination (Section 4.2). This choice will determine what algorithmic differentiation grounds can be caught by the ECHR's (initial) protective scope, and as a consequence, what algorithmic groups can find protection under the law. The Chapter's third section will examine how discriminatory actions and behaviour, and the inequalities associated thereto, can be justified (Section 4.3). The goal is to concretise the normative underpinnings that underlie the ECtHR's understanding of equality, the goals equality and non-discrimination look to achieve within society and the harms they look to guard people against. The resulting picture is then transposed onto the algorithmic environment: When evaluated through a CoE and ECtHR-inspired notion of equality, what algorithmic differentiation acts would be considered unjustified and how robust would that approach be in the face of the egalitarian harm algorithms risk imposing?

4.1 The Legal Framework

Within the ECHR, the principle of equality is given expression within Article 14 through a general prohibition of discrimination:

> *"The enjoyment of the rights and freedoms set forth in this Convention shall be secured without discrimination on any ground such as sex, race, colour, language, religion, political or other opinion, national or social origin association with a national minority, property, birth or other status."*

Article 14 of the ECHR does not enshrine a general principle of equality. It only prohibits the discrimination of people within their enjoyment of Convention rights.[363] Article 14 ECHR therefore an accessory right: it has no independent existence on its own and must be invoked in tandem with another Convention right. [364] Still, the potential and importance of Article 14 ECHR should not be understated. First, Article 14 ECHR has maintained a degree of autonomy as its application does not presuppose a violation of a substantive Convention right.[365] The prohibition to discriminate extends beyond the enjoyment of the rights and freedoms guaranteed by the Convention. The ECtHR has clarified that in order to find application, it is necessary, but also sufficient, for a case to fall "within the ambit" of one or more Convention rights.[366] Second, the ECtHR's interpretation of discrimination also applies to Article 1 of Protocol No. 12 ECHR. This protocol sought to establish a general prohibition of discrimination.[367] More specifically, Article 1 Protocol 12 guarantees the enjoyment of any right set forth by law without discrimination on any ground or any status. Additionally, the protocol stipulates that no one shall be discriminated against by any public authority on any ground.[368] However, it should be noted that many CoE Member States have not yet ratified the Protocol. Finally, the ECtHR has played a crucial role in shaping the legal interpretation of equality and non-discrimination within the CoE legal order and beyond. As the ECtHR is bound by its competences, and the boundaries of the ECHR, not every instance of differentiation or inequality can find protection under Article 14 or Article 1 Protocol

---

[363] Following the aforementioned, algorithmic differential treatment may only be scrutinized as a potential violation of the prohibition to discrimination if such would also involve a potential infringement of another ECHR right. For example, if personal data were involved, the use of which could conflict with the right to privacy.

[364] Oddný Mjöll Arnardóttir, 'Discrimination as a Magnifying Lens: Scope and Ambit under Article 14 and Protocol No. 12', in *Shaping Rights in the ECHR: The Role of the European Court of Human Rights in Determining the Scope of Human Rights*, ed. Eva Brems and Janneke Gerards (Cambridge: Cambridge University Press, 2014), 331, https://doi.org/10.1017/CBO9781107337923.020. Arnardóttir, 469.

[365] Council of Europe, 'Guide on Article 14 of the European Convention on Human Rights and on Article 1 of Protocol No. 12 to the Convention: Prohibition of Discrimination.' (Council of Europe, 31 December 2020), 7, https://www.echr.coe.int/Documents/Guide_Art_14_Art_1_Protocol_12_ENG.pdf.

[366] See for example: ECtHR (7 December 1976) Case of Kjeldsen, Busk Madsen and Pedersen v. Denmark, Application no. 5095/71; 5920/72; 5926/72, § 56. ECtHR Grand Chamber (16 March 2010) Case of Carson and Others v. The United Kingdom, Application no. 42184/05, § 63. For more information on the ambit doctrine, see also: Arnardóttir, 'Discrimination as a Magnifying Lens'.

[367] ECtHR Grand Chamber (22 December 2009) Case of Sejdic and Finci v. Bosnia and Herzegovina, § 55. Protocol No. 12 to the Convention for the Protection of Human Rights and Fundamental Freedoms; Council of Europe, Explanatory Report to the Protocol No. 12 to the Convention for the Protection of Human Rights and Fundamental Freedoms (1 November 2000); available via: https://rm.coe.int/09000016800cce48.

[368] Council of Europe, Explanatory Report to the Protocol No. 12 to the Convention for the Protection of Human Rights and Fundamental Freedoms (1 November 2000); available via: https://rm.coe.int/09000016800cce48.

12 to the ECHR.[369] Still, over the years, the ECtHR has developed and applied key concepts and doctrines regarding the fundamental scope and function of equality and non-discrimination in ways that transcend the ECHR's boundaries, substantially shaping our understanding of both notions as foundational structuring values within society.[370]

A LIVING DOCUMENT – The legal conception of equality is not merely determined by the judiciary. The ECHR is a living document, and the ECtHR must have regard to changing conditions within the Member States. Among others, the ECtHR must respond to any emerging consensus as to the standards to be achieved concerning the protection of fundamental rights.[371] This approach allows the Court to respond to challenges in a dynamic and evolutive way, interpreting rights in a practical and effective manner.[372] More specifically, the Court will "reassess 'in the light of present-day conditions"'[…] the appropriate interpretation and application of the Convention."[373] Through this approach, the ECtHR has been able to develop its understanding of equality and non-discrimination in a gradual and incremental manner.[374] At the same time, it has been said that the Court might look to uncover the moral truth underlying the ECHR, that is, to identify the moral values that act as a constraint upon state actions, rather than a commonly accepted European

[369] One obvious limitation is the ECHR's focus on the protection of civil and political rights. Hence, the ECtHR has less tools to combat discrimination and inequality in the social, economic, and cultural realm. Still, this strict separation between the civil and political sphere on the one hand, and the socio-economic and cultural sphere on the other must be nuanced. See also Section 4.3.2. The Better Placed Argument.

[370] See also: Gerards and Zuiderveen Borgesius, 'Protected Grounds and the System of Non-Discrimination Law in the Context of Algorithmic Decision-Making and Artificial Intelligence', June 2022.

[371] See for example ECtHR Grand Chamber (28 May 2002), Case of Stafford v. the United Kingdom, Application no. 46295/99, § 68. ECtHR (16 November 2014), Case of Ünal Tekeli v. Turkey, Application no. 29865/96, § 54. The legitimacy of the evolutive living document approach has been subject to debate. See also: Stefan Theil, 'Is the "Living Instrument" Approach of the European Court of Human Rights Compatible with the ECHR and International Law?', September 2017, https://doi.org/10.17863/CAM.8478.

[372] ECtHR Grand Chamber (28 May 2002), Case of Stafford v. the United Kingdom, Application no. 46295/99, § 68

[373] ECtHR Grand Chamber (28 May 2002), Case of Stafford v. the United Kingdom, Application no. 46295/99, § 69.

[374] See for instance: Janneke Gerards, 'Margin of Appreciation and Incrementalism in the Case Law of the European Court of Human Rights', *Human Rights Law Review* 18, no. 3 (1 September 2018): 495–515, https://doi.org/10.1093/hrlr/ngy017. However, it is not a given that Courts progress towards a more substantive or goal-oriented conceptualisation of equality. See for instance: Havelková, 'Judicial Scepticism of Discrimination at the ECtHR'.

standard.[375] The protection offered by the Convention should after all not be theoretical or illusory.[376]

4.2 The (Initial) Range of Protection

Article 14 ECHR and Article 1 of Protocol 12 make explicit reference to several differentiation grounds. Both clauses are phrased in an open manner as indicated by the wording that discrimination "on any ground" or "status" should be prohibited.[377] The question has nonetheless been raised whether any instance of differential treatment or impact could constitute discrimination, or whether, there are definitional and conceptual limitations to the latter notion. Over the years, the ECtHR has evolved from a fully open model towards equality and non-discrimination to one that is narrower in scope. Unfortunately, the ECtHR has not always been coherent or consistent and its case-law at times conflicting. The ECtHR approach is best described as one that is rather open-ended, but (likely) conditional to certain limitations. In the following section, I will discuss the various approaches found within ECtHR case-law and conclude with a note on the impact this has for algorithmic decision-making.[378]

An Open, Grounds-Based Approach Made Conditional?

THE UNCONDITIONAL APPROACH - The ECtHR initially advanced a clear, open-ended, and unconditional approach towards differentiation. Following the wording of the Convention, the ECtHR emphasised the illustrative and non-exhaustive nature of the grounds listed. Likewise, the

[375] G. Letsas, 'Strasbourg's Interpretive Ethic: Lessons for the International Lawyer', *European Journal of International Law* 21, no. 3 (1 August 2010): 531 and 541, https://doi.org/10.1093/ejil/chq056. See also: G. Letsas, 'The Truth in Autonomous Concepts: How To Interpret the ECHR', *European Journal of International Law* 15, no. 2 (1 April 2004): 279–305, https://doi.org/10.1093/ejil/15.2.279.

[376] See for instance: ECtHR Grand Chamber (28 May 2002), Case of Stafford v. the United Kingdom, Application no. 46295/99, § 68; ECtHR Grand Chamber (11 July 2002), Case of Christine Goodwin v. The United Kingdom (Application no. 28957/95), § 74.

[377] Arnardóttir, *Equality and Non-Discrimination under the European Convention on Human Rights*, 33.

[378] For an initial overview of the case-law, see: Gerards, 'The Discrimination Grounds of Article 14 of the European Convention on Human Rights'. See also: Arnardóttir, 'The Differences That Make a Difference'. For a more extensive application of the case-law onto algorithmic grounds, reference can be made to: Naudts, 'Criminal Profiling and Non-Discrimination: On Firm Grounds for the Digital Era?'

Court observed how the notion "other status" must be given wide meaning.[379] Under an unconditional approach, no prior conceptual or definitional constraints are placed upon the notion of discrimination. Any form of differentiation has the potential to be discriminatory and can be made subject to evaluation.[380] The Court's normative analysis therefore shifts entirely to the justification phase where it will assess whether a practice could be reasonably and objectively justified.

THE "PERSONAL CHARACTERISTIC" APPROACH – In later cases , the Court attached greater significance to the notion "other status," which acts as a sort of gate-keeper for differentiation grounds to find protection under the Convention.[381] For instance, in the case of *Kjeldsen, Busk Madsen and Pedersen*, the ECtHR noted that for differential treatment to be considered discriminatory, it must have "as its basis or reason a personal characteristic ("status") by which persons or groups of persons are distinguishable from each other."[382]

PERSONAL AS A WIDE CONCEPT - In *Clift,* the ECtHR clarified the notion "other status" in relation to the grounds expressly listed within the Convention.[383] The ECtHR first reiterated its position under *Kjeldsen, Busk Madsen and Pedersen*: "Article 14 does not prohibit all differences in treatment but only those differences based on an identifiable, objective or personal characteristic, or 'status', by which persons or groups of persons are distinguishable from one another."[384] At the same time, the notion status has "wide meaning."[385] Whereas, "a number of the specific examples [listed in Article 14 ECHR] relate to characteristics which can be said to be 'personal' in the sense that they are innate characteristics or inherently linked to the identity or the personality of the individual […] not all of the grounds listed can be thus characterised."[386] Sex, race and religion for instance can be identified as personal characteristics that are innate or inherent, whereas the characteristic property cannot.[387] The Court found it clear that "while it has

[379] ECtHR (8 June 1976) Case of Engel and Others v. The Netherlands, Application no. 5100/71; 5101/71; 5102/71; 5354/72; 5370/72, § 72. The Court confirmed this approach in Rasmussen. ECtHR (28 November 1984) Case of Rasmussen v. Denmark, Application no. 8777/79 , § 34.

[380] Gerards, 'The Discrimination Grounds of Article 14 of the European Convention on Human Rights', 104.

[381] Gerards, 104.

[382] ECtHR (7 December 1976) Case of Kjeldsen, Busk Madsen and Pedersen v. Denmark, Application no. 5095/71; 5920/72; 5926/72, § 56.

[383] ECtHR (13 July 2010, 22 November 2010, Final) Case of Clift v.The United Kingdom, Application no. 7205/07, §56.

[384] ECtHR (13 July 2010, 22 November 2010, Final) Case of Clift v.The United Kingdom, Application no. 7205/07, §55.

[385] ECtHR (13 July 2010, 22 November 2010, Final) Case of Clift v.The United Kingdom, Application no. 7205/07, §56.

[386] ECtHR (13 July 2010, 22 November 2010, Final) Case of Clift v.The United Kingdom, Application no. 7205/07, §56.

[387] ECtHR (13 July 2010, 22 November 2010, Final) Case of Clift v.The United Kingdom, Application no. 7205/07, §56.

consistently referred to the need for a distinction based on a "personal" characteristic in order to engage Article 14 […] the protection conferred by that Article is *not* limited to different treatment based on characteristics which are personal in the sense that they are innate or inherent."[388] As a consequence, the protection of non-listed grounds can extend to characteristics that are personal in the sense that they are innate or inherent, such as a person's sexual orientation and disability, as well as personal characteristics that are not, such as a person's military rank.[389] In the case of Magee, the ECtHR for instance argued that a difference between people explained in terms of the geographical location where they were arrested and detained, cannot be explained in terms of personal characteristics, such as national origin, and consequently, cannot amount to discriminatory treatment.[390]

The ECtHR nonetheless specified that "the question whether there is a difference of treatment based on a personal or identifiable characteristic in any given case is a matter to be assessed taking into consideration all of the circumstances of the case and bearing in mind that the aim of the Convention is to guarantee not rights that are theoretical or illusory but rights that are practical and effective."[391]

PERSONAL AS ANALOGUE OR SIMILAR – The ECtHR came to a different conclusion in *Peterka*.[392] The ECtHR argued that the grounds listed in Article 14 ECHR, while indicative and non-exhaustive, do have a guiding value. In order to find protection, non-listed grounds should be analogous or similar to those explicitly mentioned. The latter the ECtHR argued, reflect personality in two key ways: some characteristics reflect personal choice over one's personal identity (e.g., political opinion), other characteristics are not exercised by choice but nonetheless (inherently) tied to

---

[388] ECtHR (13 July 2010, 22 November 2010, Final) Case of Clift v. The United Kingdom, Application no. 7205/07, §59.

[389] ECtHR (13 July 2010, 22 November 2010, Final) Case of Clift v.The United Kingdom, Application no. 7205/07, §57-58.

[390] Case of Magee v the United Kingdom App no. 28135/95, § 50. First in Magee and confirmed in the Big Brother Watch case. ECtHR (20 June 2000) Case of Magee v the United Kingdom App no. 28135/95, § 50. ECtHR First Section (13 September 2018) Case of Big Brother Watch and Others v The United Kingdom App nos 58170/13, 62322/14 and 24960/15, § 516–518.

[391] Gerards rightfully observes however how the Court's position was somewhat confusing here: "the Court seemed to make clear that the notion of 'personal characteristic' is the only proper interpretation of 'other status', yet also that it may find that distinctions based on impersonal characteristics may require the Court's attention if the circumstances of the case so require." Gerards, 'The Discrimination Grounds of Article 14 of the European Convention on Human Rights', 107–9. ECtHR (13 December 2011) Case of Laduna v. Slovakia, Application no. 31827/02, § 55. ECtHR (13 July 2010, 22 November 2010, Final) Case of Clift v.The United Kingdom, Application no. 7205/07, §60.

[392] ECtHR (4 May 2010) Josef Peterka v. Czech Republic, Application no. 21990/08.

one's person (e.g., ethnicity). Characteristics that do not reflect such personal traits, convictions, or choices, and are subject to change in aleatory fashion, cannot constitute "other status" and their use cannot be considered discriminatory.[393]

PERSONAL AS INNATENESS – The ECtHR imposed an "innateness" condition onto differential treatment in the case of *Springett.* The ECtHR observed that having or not-having a right to a welfare benefit does not constitute a personal characteristic: "Unlike the principal grounds set out in Article 14, such as "sex", "race" or "colour", it is not an innate characteristic which applies from birth."[394] The ECtHR continued by stating that "unlike 'religion', 'political or other opinion' or even place of residence, the fact of having acquired a right to a benefit does not relate to a core or personal belief or choice."[395] A similar line of reasoning was followed in the case of *Čadek and others*, where the ECtHR argued that Article 14 ECHR is only protective of immutable traits, noting that a discrimination claim is manifestly ill-founded in cases where "different treatment […] was not based on any personal choice in so far as this choice should be respected as elements of someone's personality, such as religion, political opinion, sexual orientation and gender identity, or on grounds of personal features in respect of which no choice at all can be made, such as sex, race, disability and age."[396]

---

[393] In the words of the Court: "De l'avis de la Cour, si la liste des motifs de discrimination dressée à l'article 14 est, certes, indicative, l'on ne saurait affirmer pour autant que cette disposition interdit toute discrimination, quel que soit le critère qui la fonde." Unfortunately, the Court also included sexual orientation and identity as "choice" grounds. The Court noted: "Il peut s'agir des raisons tenant à des choix personnels reflétant des traits de la personnalité de chacun, comme la religion, les opinions politiques, *l'orientation sexuelle et l'identité sexuelle*, ou des motifs tenant à des caractéristiques personnelles au sujet desquelles on ne peut exercer aucun choix, comme le sexe, la race, le handicap et l'âge." As a consequence of this reasoning, the ECtHR found that differences in duration of contract or differences in employment status (*in casu,* the difference between employee or independent) did not reflect any personal trait, conviction or choice and can therefore not be covered by Article 14. Interestingly, the ECtHR did refer to the grounds listed article 21 of the EU Charter of Fundamental Rights as a means to further interpret the scope of Article 14 ECHR. ECtHR (4 May 2010) Josef Peterka v. Czech Republic, Application no. 21990/08. See also Gerards, 'The Discrimination Grounds of Article 14 of the European Convention on Human Rights', 108.

[394] ECHR (27 April 2010) Springett and Others v United Kingdom, Application No. 34726/04, 14287/05 and 34702/05 (Admissibility), § 7.

[395] ECHR (27 April 2010) Springett and Others v United Kingdom, Application No. 34726/04, 14287/05 and 34702/05 (Admissibility), § 7.

[396] ECtHR (22 November 2012; Final 29 April 2013) Case of Čadek and Others v. The Czech Republic, Applications nos. 31933/08, 60084/08, 6185/09, 46696/09, 52792/09, 53518/09, 10185/10, 42151/10, 3167/11 and 20939/11, § 94. In this case, the ECtHR had to consider if differential treatment based on whether someone is an original restitution claimant or bought a restitution claim is a relevant ground under Article 14 ECHR.

SETTLED APPROACH? – In their recent guide on Article 14 ECHR, the CoE seems to have confirmed a departure from a fully-open approach: "Article 14 does not prohibit all differences in treatment but only those based on an identifiably, objective or personal characteristic or 'status', by which persons or groups of persons are distinguishable from one another."[397] Still, the notion "other status" should be given wide meaning and not be limited to personal characteristics that are innate or inherent.[398] While this does not impose a condition of immutability, the condition that differentiation grounds should be personal and identifiable appears maintained.[399]

Unfortunately, even if the "other status" notion must be given broad meaning, little guidance has been offered as to when a ground will be considered personal or not.[400] Because the protective

---

[397] In recent cases, the description has not always included the notion "personal": "only differences in treatment based on an identifiable characteristic, or "status", are capable of amounting to discrimination within the meaning of Article 14." See ECtHR (4 May 2021) Case of Jurčić v. Croatia, Application no. 54711/15, § 63. Additionally, the ECtHR has specified that, once a ground has been accepted as another status, that ground need not be equated with those explicitly mentioned. However, the latter is more closely related to the idea that certain grounds trigger a higher level of scrutiny. ECtHR (15 September 2016, Final 30 January 2017) Case of British Gurtkha Welfare Society and Others v. the United Kingdom, § 88. See also: ECtHR (25 July 2017, Final 25 October 2017) Case of Carvalho Pinto de Sousa Morais v. Portugal, Application no. 17484/15, §45.

[398] Council of Europe, 'Guide on Article 14 of the European Convention on Human Rights and on Article 1 of Protocol No. 12 to the Convention: Prohibition of Discrimination.', 34. A recent example involves the *Molla Sali* case, where the ECtHR moreover specified that the "Convention covers instances in which an individual is treated less favourably on the basis of another person's status or protected characteristics. ECtHR Grand Chamber (2018) Molla Sali v. Greece, Application no. 20452/14., § 134. Arnardóttir, 'Vulnerability under Article 14 of the European Convention on Human Rights', 154. See also: ECtHR Grand Chamber (16 March 2010) Case of Carson and Others v. The United Kingdom, Application no. 42184/05, § 70; ECtHR (13 July 2010, 22 November 2010, Final) Case of Clift v.The United Kingdom, Application no. 7205/07, §56.

[399] Over the years, the Court has recognized several grounds as being a protected differentiation falling under the "other status" rubric. Such examples include, gender identity, sexual orientation, health and disability, parental and marital status, immigration and employment status. For a full overview, see Council of Europe, 'Guide on Article 14 of the European Convention on Human Rights and on Article 1 of Protocol No. 12 to the Convention: Prohibition of Discrimination.' Examples highlighted by the Court of grounds not considered personal (or not falling within other status), include: holding fishing rights in different areas, being sent on different military missions, difference in legal status concerning the restitution of taxes, the distinction between smaller and larger unions and the difference between commercial advertising and advertisements forwarding certain ideals. Council of Europe, 41. As aforementioned, geographic location has also been excluded from protection. This was first established in Magee and later confirmed in the Big Brother Watch case. ECtHR (20 June 2000) Case of Magee v the United Kingdom App no. 28135/95, § 50. ECtHR First Section (13 September 2018) Case of Big Brother Watch and Others v The United Kingdom App nos 58170/13, 62322/14 and 24960/15, § 516–518.

[400] Some authors argue that much confusion stems from the fact that there is no concrete rationale or theory underlying the development of Article 14, but instead the Court's approach has been developed around a plurality of conceptions.

net of Article 14 ECHR is therefore cast rather wide, it may prove difficult to ascertain beforehand whether or not a ground will be excluded on the basis of the aforementioned case-law - unless the Court has a precedent for its inclusion or exclusion.

COMPLEX DIFFERENTIATION – The requirement that differentiation grounds must be identifiable in order to receive protection could be limiting in the face of structural, systemic, or intersectional discriminatory harms. These harms cannot always be reduced to one single, identifiable ground or cause. Instead, a person's disadvantageous position is the result of a variety of multiple interacting and intersecting factors. Furthermore, the ECtHR has not always recognised the cumulative effects that discrimination can have where it is based on various "differentiation grounds."[401] For instance, in the case of *BS v. Spain*, the Court observed how the domestic courts failed to take into account the particular vulnerability of the applicant "inherent in her position as an African woman working as a prostitute."[402] Whereas the notion of intersectionality could have made the difficulties faced by those multiply burdened more explicit, the interaction of multiple differentiation grounds and the complex nature of discrimination was implicitly recognised by way of the notion of vulnerability.[403] In a similar vein, in the case of *Carvalho Pinto de Sousa Morais*, the ECtHR condemned the domestic court's reliance on the generalised assumption that sexuality

---

This "plurality" will be discussed in detail under the ECtHR's justification. Anna Nilsson, 'Same, Same but Different: Proportionality Assessments and Equality Norms', *Oslo Law Review* 7, no. 3 (17 December 2020): 144, https://doi.org/10.18261/ISSN.2387-3299-2020-03-01; Gerards, 'The Discrimination Grounds of Article 14 of the European Convention on Human Rights', 113; Arnardóttir, *Equality and Non-Discrimination under the European Convention on Human Rights*, 6–7; Arnardóttir, 'The Differences That Make a Difference', 669.

[401] Lisa Maria Weinberger, 'Kurt v. Austria: A Missed Chance to Tackle Intersectional Discrimination and Gender-Based Stereotyping in Domestic Violence Cases', *Strasbourg Observers* (blog), 8 August 2021, https://strasbourgobservers.com/2021/08/18/kurt-v-austria-a-missed-chance-to-tackle-intersectional-discrimination-and-gender-based-stereotyping-in-domestic-violence-cases/; Senem Gurol, 'Challenging Gender Stereotyping before the ECtHR: Case of Carvalho Pinto v. Portugal', *EJIL: Talk!* (blog), 21 September 2017, https://www.ejiltalk.org/challenging-gender-stereotyping-before-the-ecthr-case-of-carvalho-pinto-v-portugal/.
Lourdes Peroni, 'Age and Gender Discrimination: Laudable Anti-Stereotyping Reasoning in Carvalho Pinto V. Portugal', *Strasbourg Observers* (blog), 28 September 2017, https://strasbourgobservers.com/2017/09/28/age-and-gender-discrimination-laudable-anti-stereotyping-reasoning-in-carvalho-pinto-v-portugal/. ECtHR Grand Chamber (15 June 2021) Case of Kurt v. Austria (Application no. 62903/15)

[402] ECtHR (24 July 2012, Final 24 October 2012) Case of B.S. v. Spain, Application no. 47159/08, § 62.

[403] See also Chapter 2, Section 2.3 "Equality and Intersectional Differentiation." ECtHR (24 July 2012, Final 24 October 2012) Case of B.S. v. Spain, Application no. 47159/08, § 56 and 57. The notion of intersectional discrimination was brought to the attention by the case's third-party interveners. Arnardóttir, 'Vulnerability under Article 14 of the European Convention on Human Rights'.

would not be important to older women and mothers as it would be for a younger woman.[404] In its ruling, the ECtHR relied upon an anti-stereotyping approach without explicitly recognising the intersectional nature of the stereotype at hand.[405] It is therefore not entirely clear whether the ECtHR would approach equality as a single- or multi-axis paradigm.[406] Taking into consideration the ECtHR's more open approach towards Article 14 ECHR, it is likely that complex forms of discrimination, including intersectional discrimination, are not precluded *a priori* from finding protection under the Court's equality conception.[407]

ALGORITHMS AND DIFFERENTIATION GROUNDS[408] **–** What can we learn from the above-mentioned observations? First, the ECtHR's approach towards equality is grounds-based. Second, the ECtHR's conceptualisation of equality may allow for the protection of more complex forms of (multi-ground) differentiation. Third, this protection is made conditional: differential treatment, and the inequalities that flow thereout, must be based on an objective, identifiable and personal characteristic. As a consequence, the Court's interpretation of equality and non-discrimination does not *a priori* preclude the protection of algorithmic inequalities. as long as the underlying differentiation ground, which could be complex, can be considered objective, identifiable, and personal.

---

[404] ECtHR (25 July 2017, Final 25 October 2017) Case of Carvalho Pinto de Sousa Morais v. Portugal, Application no. 17484/15, § 52-56.

[405] In this regard, scholars have observed how the ECtHR downplayed the importance of the age dimension, focusing primarily on sexuality. See: Gurol, 'Challenging Gender Stereotyping before the ECtHR'; Peroni, 'Age and Gender Discrimination: Laudable Anti-Stereotyping Reasoning in Carvalho Pinto V. Portugal'.

[406] For example, in the case of *Carvalho Pinto* judges Ravarani and Bošnjak clarified in a joint dissenting opinion that the Court failed to illustrate the applicant's complaint regarding age discrimination, while also failing to analyse the claim of gender discrimination. The requirement to have both forms of discrimination illustrated separately is indicative of single-axis reasoning. An intersectional model would explicitly reject such a view and would instead highlight the inseparable nature of the various differentiation grounds at play. ECtHR (25 July 2017, Final 25 October 2017) Case of Carvalho Pinto de Sousa Morais v. Portugal, Application no. 17484/15, p. 31 – 43.

[407] In their guide on Article 14 ECHR seem to confirm this viewpoint: "The Court has also examined situations of discrimination that took place on the basis of several grounds operating separately or interacting with each other at the same time. Both Article 14 of the Convention and Article 1 of Protocol No. 12 prohibit discrimination on a large number of grounds, making a claim on more than one ground theoretically possible. Furthermore, the non-exhaustive list of grounds of discrimination contained in Article 14 allows the Court to extend and include grounds not expressly mentioned therein." Council of Europe, 'Guide on Article 14 of the European Convention on Human Rights and on Article 1 of Protocol No. 12 to the Convention: Prohibition of Discrimination.', 9–10.

[408] This paragraph is an adaptation of: Naudts, 'Criminal Profiling and Non-Discrimination: On Firm Grounds for the Digital Era?'

Conceptual difficulties do start to emerge however when one tries to define a differentiation ground as objective, identifiable and personal. Without this clarity, we do not know whether the Court's understanding of equality can effectively capture algorithmic forms of differentiation. For instance, in relation to the identifiability criterion, algorithms pose a challenge as they can function in ways that are opaque and dynamic.[409]

While challenges of complexity, opacity and explicability could be overcome to some extent, the core limitation that remains within the Court's definition of discrimination concerns the requirement that differential treatment must be based upon a personal characteristic. To what extent are algorithmic differentiation grounds personal? Sure, an algorithmic differentiation ground is applied *onto people*, and often the ground will be derived from an aggregate of information *about or relating to people*, but does this render a differentiation ground a personal characteristic? Should the notion "objective, identifiable or personal characteristic" be equated with information *related to a person*?[410] For instance, could average school results, be considered a status that reflects a personal characteristic? What about the evaluation of people in accordance

[409] See also: Wachter, 'The Theory of Artificial Immutability'.

[410] The notion of an identifiable personal characteristic under ECtHR equality law might not correspond to the notion of "personal data" under EU data protection law. Personal data concerns any information related to an *identified* or *identifiable* natural person. In the definition of personal data, the notion "identifiable" pertains to the natural person that is captured by a given piece of information. In the case of the ECtHR however, the notion "identifiable" pertains to the characteristic as being identifiable, rather than the persons involved. Furthermore, when I say that the notion of an identifiably personal characteristic under the ECtHR could be interpreted as meaning a characteristic "related to a person" this understanding would be broader than the notion of personal data: it does not impose the requirement that the characteristic involved makes it possible for the persons involved to be identified or identifiable. In my use, the idea that a characteristic relates to a person would simply imply that that characteristic can find application, and as a consequence, affect persons. The notion thus covers information pertaining to algorithmic groups or aggregates. Even if the ECtHR ascribes wide meaning to the notion status, both the concept of "personal data" or "information related to a person" could be too broad. If the Court wants to limit the ECtHR's protective scope by placing boundaries on the number of differentiation grounds that can constitute discrimination, reliance on either notion would expand the scope protection the ECtHR is willing to cover. For example, the differentiation ground "geographic location", which has been denied protection, is both personal data and data related to a person. Vedder, 'KDD'; Linnet Taylor, Luciano Floridi, and Bart van der Sloot, eds., *Group Privacy* (Cham: Springer International Publishing, 2017), https://doi.org/10.1007/978-3-319-46608-8; Alessandro Mantelero, 'From Group Privacy to Collective Privacy: Towards a New Dimension of Privacy and Data Protection in the Big Data Era', in *Group Privacy: New Challenges of Data Technologies*, ed. Linnet Taylor, Luciano Floridi, and Bart van der Sloot (Cham: Springer International Publishing, 2017), 139–58, https://doi.org/10.1007/978-3-319-46608-8_8; Lanah Kammourieh et al., 'Group Privacy in the Age of Big Data', in *Group Privacy*, ed. Linnet Taylor, Luciano Floridi, and Bart van der Sloot (Cham: Springer International Publishing, 2017), 37–66, https://doi.org/10.1007/978-3-319-46608-8_3; Mittelstadt, 'From Individual to Group Privacy in Big Data Analytics'.

with non-tangible traits, such as their viewing behaviour? Even in the case certain (individual) traits, such as education, political or philosophical beliefs, could fall under other status definition, a profile might still comprise other traits that do not fit the latter category. Could the combined profile still be considered personal, especially when taking into account that the differentiation ground could constitute generalised, rather than individual or personal information?[411] Finally, there are certain traits that clearly have great importance within the algorithmic environment, but that the ECtHR has already deemed as non-personal, such as geographic location.[412] If one wishes to ascertain to what extent algorithmic egalitarian harms can find redress under the ECtHR's conceptualisation of equality, these conceptual challenges would need to be resolved.[413]

4.3 The Justification of Differentiation

Even if the ECtHR's conceptualisation of equality has the potential to capture many forms of (algorithmic) differential treatment, not every algorithmic differentiation ground made subject to evaluation will actually amount to a violation of the general principle of equality or prohibition of discrimination. The Court must still assess whether the differentiating act can be objectively and reasonably justified. This process of evaluation gives further shape to the ECtHR's conceptualisation of equality.[414] In the following paragraphs, two key elements that shape this conceptualisation will be discussed.

On the one hand, the Court has advanced a standardised review procedure for discrimination cases. On the other hand, over the years, the ECtHR has developed certain concepts and doctrines (or contextual factors) to further specify the standard evaluative approach. More specifically, for certain forms of differential treatment, the Court will heighten its level of scrutiny, either because the States' margin of appreciation is lower, or because weighty reasons must be adduced where specific differentiation grounds are used. These contextual factors clarify those inequalities the

[411] Naudts, 'Criminal Profiling and Non-Discrimination: On Firm Grounds for the Digital Era?', 81.

[412] ECtHR (20 June 2000) Case of Magee v the United Kingdom App no. 28135/95, § 50. ECtHR (13 September 2018) Case of Big Brother Watch and Others v The United Kingdom App nos 58170/13, 62322/14 and 24960/15, § 516–518.

[413] Of course, in the case a personal characteristic should also be innate, inherent, or immutable, the protective scope becomes even more narrow. Furthermore, this requirement too would lead to conceptual difficulties when juxtaposed onto the algorithmic environment When, and under what conditions, could algorithmic differentiation grounds constitute an innate or inherent personal trait? See also Chapter 3, Section 4.3 "Immutability": Naudts, 'Criminal Profiling and Non-Discrimination: On Firm Grounds for the Digital Era?', 81. See also on this problem: Wachter, 'The Theory of Artificial Immutability'.

[414] Or more precisely, the Court's conceptualisation of equality within its competences.

ECtHR can address, as well as those inequalities the ECtHR deems particularly nefarious. Taken together, the standard review procedure and the elements that raise (or lower) the Court's level of review, can be used to further uncover the ECtHR's conceptualisation of equality, and how (in)equalities and discriminatory acts ought to be balanced against competing interests.

4.3.1 The Review Procedure

Once the ECtHR has found that a differentiation ground falls within the scope of Article 14 ECHR, it will decide whether differential treatment based upon that ground can be justified. In order to do so, the ECtHR has developed a three-step test. First, in order for an issue to arise, the difference of treatment must affect persons in analogous or relevantly similar situations (or vice versa, similar treatment must affect persons in relevantly different situations).[415] In a second phase, the Court will investigate whether this treatment was objectively justified.[416] The ECtHR will investigate, i) whether the measure pursued a legitimate aim, and ii) whether there was a "reasonable relationship" of proportionality between the means employed and the aims a measure sought to realise.[417]

Comparability

Under the first step, the ECtHR examines whether or not there was a difference in treatment of persons in analogous or relevantly similar situations (or the reverse, was there similar treatment of people in different situations). This review is typically made through a comparative exercise: is there a similar group or person that was in a comparable (or different) situation, but nonetheless benefited from different (or similar) treatment? The need to demonstrate an "analogous position" does not require that the comparator groups be identical. The question whether or not person or groups find themselves in comparable situations, necessitates a specific and context-dependent

[415] See for example ECtHR Grand Chamber (19 December 2018) Case of Molla Sali v Greece, Application no. 20452/14; §135. uncil of Europe, 'Guide on Article 14 of the European Convention on Human Rights and on Article 1 of Protocol No. 12 to the Convention: Prohibition of Discrimination.', 16.

[416] See for example ECtHR Grand Chamber (19 December 2018) Case of Molla Sali v Greece, Application no. 20452/14; §135. For an early account, see the ECHR (1968), Case "Relating to certain aspects of the laws on the use of languages in education in Belgium" v Belgium, Application no. 1474/62; 1677/62; 1691/62; 1769/63; 1994/63; 2126/64.

[416] See for example ECtHR Grand Chamber (19 December 2018) Case of Molla Sali v Greece, Application no. 20452/14; §135

[417] See for example ECtHR Grand Chamber (19 December 2018) Case of Molla Sali v Greece, Application no. 20452/14; §135

analysis, taking into consideration the subject-matter and purpose of the measure that imposes the distinction in question.[418]

The need for comparison is seen as an exponent of formal equality. The postulate that like situations should be treated alike assumes that most inequalities can be identified through a comparative exercise.[419] The Court has shown inconsistent use of the comparability requirement however, and sometimes the comparator approach has been discarded.[420] Moreover, in certain cases, judges appear divided as to whether they should engage into a comparative exercise, or whether they should instead focus on the concrete disadvantage or harms experienced by a given person or group.[421] Proponents of a formal comparative approach tend to invoke legitimacy as a

---

[418] ECtHR Grand Chamber (5 September 2017), Case of Fabian v. Hungary, Application no. 78117/13, § 121.

[419] Katarina Frostell, 'Gender-Based Violence Triggers Differential Treatment in Housing Benefit Case', *Strasbourg Observers* (blog), 2 December 2019, https://strasbourg.weichie.dev/2019/12/02/gender-based-violence-triggers-differential-treatment-in-housing-benefit-case/. Arnardóttir, 'Vulnerability under Article 14 of the European Convention on Human Rights', 154. O. M. Arnardottir, 'The Differences That Make a Difference: Recent Developments on the Discrimination Grounds and the Margin of Appreciation under Article 14 of the European Convention on Human Rights', *Human Rights Law Review* 14, no. 4 (1 December 2014): 657, https://doi.org/10.1093/hrlr/ngu025. Arnardóttir, *Equality and Non-Discrimination under the European Convention on Human Rights*, 23; Alexandra Timmer and Lourdes Peroni, 'Bah v UK: On Immigration, Discrimination and Worrisome Reasoning', *Strasbourg Observers* (blog), 12 October 2011, https://strasbourgobservers.com/2011/10/12/bah-v-uk-on-immigration-discrimination-and-worrisome-reasoning/. Timmer and Peroni reference the work of Goldberg for a more critical investigation of the comparator approach. Suzanne B Goldberg, 'Discrimination by Comparison', *The Yale Law Journal*, 2011, 85. See also: Peroni, 'Age and Gender Discrimination: Laudable Anti-Stereotyping Reasoning in Carvalho Pinto V. Portugal'.

[420] See on this inconsistent application of the "comparator approach" the analysis made by Kim, who also analyses the approach in the light of the ECtHR's introduction of "vulnerability" as a criterion. So Yeon Kim, 'Les Vulnérables : Evaluating the Vulnerability Criterion in Article 14 Cases by the European Court of Human Rights', *Legal Studies* 41, no. 4 (December 2021): 623, https://doi.org/10.1017/lst.2021.21.

[421] One case where the comparability approach was discussed concerns the aforementioned Carvalho Pinto de Sousa Morais v. Portugal. In this case a woman had sought damages from a hospital after having had a failed gynaecological surgery. The Portuguese Supreme Court had reduced the claim for damages. The domestic court did so (among others) upon the basis of stereotypical assumptions concerning female sexuality. The Court noted how "at the time of the operation the plaintiff was already 50 years old and had two children, that is, an age when sex is not as important as in younger years, its significance diminishing with age." In a concurring opinion, judge Yudkivska, noted that a comparison was not needed as the Supreme Court's use of prejudicial stereotypes affected the judicial assessment of evidence. This alone should have been sufficient to find a violation of Article 14 ECHR. While Judge Yudkiviska acknowledged how discrimination may be easier to identify where there is an accumulation of comparable cases, she also quoted Timmer, noting that the wrongful nature of stereotyping is not comparative in nature, in the sense that they do not derive from a comparison another group treated differently and better. Hence there was no need to enter into a comparative exercise to find discrimination. Judge Motok too concurred with the majority: once it has been established that a stereotype is harmful, the stereotype should be contested. Here, we should not use a comparator, but instead focus on

core argument against the abandonment of the comparison requirement. Under this view, judges should not enter into the realm of politics. The standardised review-procedure is seen as a safeguard against judicial activism.[422] On other hand, judges in favour of alternative evaluative mechanisms, tend to view equality as a substantive, goal-oriented principle. With a purpose to redress disadvantage and protect vulnerable groups. Whereas the latter might not result in an abandonment of the comparative approach altogether, proponents of substantive equality would nonetheless acknowledge that the approach does not fit certain types of injustice. Under this approach, certain discriminatory and egalitarian harms can stand on their own.[423] For example, if

---

the disadvantage. This means to prove a stereotype was harmful to the group to whom the applicant belongs and that the stereotype was used as a basis for differentiation. In their dissenting opinion though, judges Ravani and Bošnjak, condemned the abandonment of the comparative approach, however. They argued that there should have been a comparison between women and men suffering from physical harm and seeking compensation for non-pecuniary damage. This had to be ascertained by an analysis of the facts. Their end-conclusion was that the Court failed to apply the well-established three-step methodology without providing any reasons to do so and failed to analyse the existence of age and sex discrimination properly. They observe: "A judge who intends to deliver a message on a legal issue of general importance should wait for the right case to do so; otherwise he or she simply engages in politics. It is no more and no less than a question of legitimacy. ECtHR (25 July 2017, Final 25 October 2017) Case of Carvalho Pinto de Sousa Morais v. Portugal, Application no. 17484/15, § 52. Concurring Opinion of Judge Yudkivska, ECtHR (25 July 2017, Final 25 October 2017) Case of Carvalho Pinto de Sousa Morais v. Portugal, Application no. 17484/15, at page 22 and 24. In reference to: Alexandra Timmer, 'Judging Stereotypes: What the European Court of Human Rights Can Borrow from American and Canadian Equal Protection Law', *The American Journal of Comparative Law* 63, no. 1 (1 January 2015): 252, https://doi.org/10.5131/AJCL.2015.0007. Concurring Opinion of Judge Motoc in ECtHR (25 July 2017, Final 25 October 2017) Case of Carvalho Pinto de Sousa Morais v. Portugal, Application no. 17484/15, § 18. Joint Dissenting Opinion of Judges Ravani and Bošnjak in ECtHR (25 July 2017, Final 25 October 2017) Case of Carvalho Pinto de Sousa Morais v. Portugal, Application no. 17484/15, § 38.

[422] In J.D. and A. v. the United Kingdom, the ECtHR found a partial violation of Article 14. More specifically, they found that the UK's housing benefit scheme took insufficient account of the particularly vulnerable position of women subject to domestic violence. The majority ruling did not apply a strict comparability test, but instead looked at the disadvantage experienced, emphasising the harm and prejudicial effect of a given measure on a vulnerable group. However, this shift away from the comparator approach was criticized in the partially descending opinion by judges Wojtzeck and Pejchal. They too seemed to argue that an abondement of the formal approach presents a risk of "undue interference in the sphere of political choices." Because every legislation must differentiate, the art of legislation, they argue, boils down to the "art of wise differentiation." Not having performed a comparative exercise, they argued, the Court became an arbiter on how the government's resources were best allocated, and this should be something a judge should be cautious of. Joint Partially Dissenting Opinion of Judges Wojtyczek and Pejchal, ECtHR First Section (24 October 2019, final 24 February 2020), Case of J.D. and A. v. The United Kingdom, Application nos. 32949/17 and 34614/17, § 4, 9 and 11. See also: Frostell, 'Gender-Based Violence Triggers Differential Treatment in Housing Benefit Case'. See also footnote above and Joint Dissenting Opinion of Judges Ravani and Bošnjak in ECtHR (25 July 2017, Final 25 October 2017) Case of Carvalho Pinto de Sousa Morais v. Portugal, Application no. 17484/15.

[423] See for instance: A. Timmer, 'Toward an Anti-Stereotyping Approach for the European Court of Human Rights', *Human Rights Law Review* 11, no. 4 (1 December 2011): 724, https://doi.org/10.1093/hrlr/ngr036. Timmer and

the wrong of stereotyped and generalised decision-making lies in the fact that these processes misrepresent or misrecognise a given person or group, this disadvantage is said to exist independently. In other words, the existence of stereotyped reasoning in itself should suffice to bring forward an equality claim.[424] The assumption that inequalities can only be identified through a comparative exercise has therefore also found criticism among substantive non-discrimination scholars.[425] They have also demonstrated how in certain discrimination cases, such as cases of complex or intersectional discrimination, a person or group might be as uniquely situated that a comparable group cannot reasonably be found.[426]

Though the comparative test is claimed to be apolitical in nature, its performance is far from neutral. The comparison, Cartibia notes, is not mechanical, but necessitates the identification of a specific feature of relevance, which then acts as the main point of reference for the comparison.[427] This characteristic will define whether or not two given situations are alike. The feature of relevance, or choice of comparator, directly governs the outcome of the case.[428] Furthermore, in the case a discrimination claim is dismissed because no comparable group can be found, the

---

Peroni, 'Bah v UK: On Immigration, Discrimination and Worrisome Reasoning'. On likeness and comparability, see also: Charilaos Nikolaidis, 'Rethinking Likeness and Comparability in Equality Claims Brought before the European Court of Human Rights', *Public Law*, July 2020, 448–67.

[424] See also: Timmer, 'Toward an Anti-Stereotyping Approach for the European Court of Human Rights', 724.

[425] See: Suzanne Goldberg, 'Discrimination by Comparison', *Yale L.J.* 120 (1 January 2011): 728.

[426] See *inter alia:* Timmer, 'Toward an Anti-Stereotyping Approach for the European Court of Human Rights', 724; Evelyn Ellis and Philippa Watson, 'Key Concepts in EU Anti-Discrimination Law', in *EU Anti-Discrimination Law*, 2nd ed. (Oxford: Oxford University Press, 2012), 156, https://doi.org/10.1093/acprof:oso/9780199698462.003.0004; Gauthier De Beco, 'Harnessing the Full Potential of Intersectionality Theory in International Human Rights Law : Lessons from Disabled Children's Right to Education', in *Intersectionality and Human Rights Law*, ed. Shreya Atrey and Peter Dunne, 1st ed. (Oxford: Hart Publishing, 2020), 41, http://www.bloomsburycollections.com/book/intersectionality-and-human-rights-law/ch2-harnessing-the-full-potential-of-intersectionality-theory-in-international-human-rights-law/.

[427] M. Cartabia, 'The European Court of Human Rights: Judging Nondiscrimination', *International Journal of Constitutional Law* 9, no. 3–4 (1 October 2011): 812, https://doi.org/10.1093/icon/mor043.

[428] Cartabia, 812. See also: Kim, 'Les Vulnérables', 624. Joint Dissenting Opinion of Judges Ravani and Bošnjak in ECtHR (25 July 2017, Final 25 October 2017) Case of Carvalho Pinto de Sousa Morais v. Portugal, Application no. 17484/15, § 9. Cartabia gives the example of the Oršuš v Croatia case. This case concerned the organisation of the Croatian education system whereby Roma children were often assigned to separate, special-needs classes. According to Cartabia the conclusion as to whether such a system is discriminatory might very well differ on the basis of the point of comparison: 'Who are the "like" students to be compared with in this case? If the linguistic skills were the relevant feature to be taken into account, then the result would be a finding of non-violation; if, instead, the question of race were to prevail in the analysis, then the Court would reach the opposite result.' See also ECtHR Grand Chamber (16 March 2010) Case of Oršuš and Others v. Croatia, Application no. 15766/03.

comparative approach is liable to maintain a status-quo. This position too is value-laden and could thus be critiqued from a legitimacy perspective.

The choice for a comparative approach could also prove too lenient when used as a mechanism to govern algorithmically guided decisions. One of the core functions that algorithms perform is to group people in specific categories and to find a basis upon which they can be differentiated. The purpose of algorithmic learning techniques is to find those traits that tie people together so that it can be said that they are either in analogous situations or not. Hence, they will always provide a "relevant point of reference" for the decision-maker to legitimise the difference in treatment they impose. This poses a challenge. On the one hand, to adopt a procedural comparability approach could result in the default acceptance of every algorithmically guided or imposed form of differentiation, and egalitarian harm following thereout. On the other hand, in case the comparative approach is abandoned, judges are given the opportunity to scrutinise the particular functioning and supposed rationality of algorithms, and in so doing, they would actively enter into the algorithmic governance debate. Given the omnipresence of algorithmic decisions, both approaches could be perceived as political and illegitimate. Perhaps then a political debate is first needed concerning the value and function a comparative approach can have as an evaluative mechanism within the algorithmic environment.

Legitimate Aim

Differential treatment does not constitute discrimination where it can be objectively and reasonably justified. A differential measure can only be justified where it pursues a legitimate aim. The intensity of the ECtHR's review can be impacted by the aim pursued.[429]

An aim can be expressly recognised as legitimate within the ECtHR treaty.[430] Otherwise, the aim pursued should be compatible with the Convention.[431] It is not sufficient for differential treatment to pursue a legitimate aim. A link must also be established between the aim pursued and the

[429] For instance, the ECtHR may apply a higher degree of scrutiny where the aim is based upon, or looks to maintain, socio-relational structures of inequality. Where the aim relates to a State's socio-economic or distributive policies, the Court may apply a lower level of scrutiny.

[430] Nilsson, 'Same, Same but Different', 129. For example, Articles 8-11 ECHR are non-absolute, and they reference some goals which can justify an interference.

[431] Nilsson, 129.

impugned measure.[432] For example, it cannot be said that a measure pursues a legitimate aim when the goal in question can be preserved and consolidated without the measure being in place.[433] Similarly, a differentiating measure cannot simply be justified by making reference to a particular aim.[434] Finally, if a measure imposes differentiating treatment, the aim pursued might only be considered legitimate where appropriate and effective safeguards accompany each stage of that measure's implementation, especially where the alleged aim requires for a given group to be disproportionately affected or excluded.[435]

Over the years, the ECtHR has recognised a variety of aims pursued by differential measures as legitimate and it generally has been rather accepting of the motivations put forward by decision-

[432] See also: Council of Europe, 'Guide on Article 14 of the European Convention on Human Rights and on Article 1 of Protocol No. 12 to the Convention: Prohibition of Discrimination.', 18. In reference to: ECtHR Fourth Section (16 February 2005 final), Case of Ünal Tekeli v. Turkey, Application no. 29865/96, § 65 -66.

[433] For instance, in the Case of Ünal Tekeli v. Turkey, the ECtHR had to consider whether the need to preserve family unity is a legitimate aim for the imposition of a joint family need: "The Court observes […] it is perfectly conceivable that family unity will be preserved and consolidated where a married couple chooses not to bear a joint family name. Observation of the systems applicable in Europe supports this finding. The Government have not shown in the present case that concrete or substantial hardship for married partners and/or third parties or detriment to the public interest would be likely to flow from the lack of reflection of family unity through a joint family name. In these circumstances the Court considers that the obligation on married women, in the name of family unity, to bear their husband's surname […] has no objective and reasonable justification." In reference to: ECtHR Fourth Section (16 February 2005 final), Case of Ünal Tekeli v. Turkey, Application no. 29865/96, § 66.

[434] See also: Janneke Gerards, *Judicial Review in Equal Treatment Cases* (Brill Nijhoff, 2005), 137, http://brill.com/view/title/11318.

[435] Such would be the case for instance, where the aim requires the segregation of given group. This was the alleged aim in the case of Oršuš and Others v. Croatia. The Court observed that the "temporary placement of children in a separate class on the ground that they lack an adequate command of the language is not, as such, automatically contrary to Article 14 of the Convention. It might be said that in certain circumstances such placement would pursue the legitimate aim of adapting the education system to the specific needs of the children. However, when such a measure disproportionately or even, as in the present case, exclusively, affects members of a specific ethnic group, then appropriate safeguards have to be put in place." ECtHR Grand Chamber (16 March 2010) Case of Oršuš and Others v. Croatia, Application no. 15766/03, § 157. Whether or not a measure is capable of achieving the aims set out, is not always addressed, however. The ECtHR may directly opt to question the proportionality of a given measure instead. For example, in Molla Sali, the ECtHR expressed doubts whether a given measure was suited to achieve a particular aim. However, the Court found it unnecessary to take a firm stance on this point considering that "the impugned measure was in any event not proportionate to the aim pursued." ECtHR Grand Chamber (19 December 2018) Case of Molla Sali v. Greece, Application no. 20452/14, § 143. Council of Europe, 'Guide on Article 14 of the European Convention on Human Rights and on Article 1 of Protocol No. 12 to the Convention: Prohibition of Discrimination.', 19.

makers even if they are broad, general or abstract.[436] The aims pursued can relate to distributional policies, such as the preservation of a given economic and social(welfare) system[437], the maintenance of economic stability,[438] or adapting the education system to the needs of children.[439] Aims can also be more social in nature, such as the need to maintain a certain way of "living together" within society.[440] The preservation of legal certainty has also been considered a legitimate aim.[441]

Despite their leniency, on several occasions the aim pursued was declared illegitimate. Of particular note is the case of *Konstantin Markin v. Russia,* which concerned the exclusion of male military personnel from being entitled to parental leave*.* The ECtHR argued that "references to traditions, general assumptions or prevailing social attitudes [such as traditions that derive from the man's primordial role and the woman's secondary role in the family are] insufficient justification [to impose] a difference in treatment on grounds of sex."[442] A reference to the traditional distribution of gender roles in society cannot justify differential treatment: "gender stereotypes, such as the perception of women as primary child-carers and men as primary breadwinners, cannot, by themselves, be considered to amount to sufficient justification for a difference in treatment, any more than similar stereotypes based on race, origin, colour or sexual orientation."[443] The pursuit of these objectives would furthermore be detrimental to both the position of women and men. By reinforcing gender stereotypes, the difference in treatment would

---

[436] In light of the Court's general tendency to accept very broad aims, Gerards has argued for a more critical approach, where the review should not only investigate the aims that were expressly declared, but also examine other and more covert goals decision-makers might pursue. Gerards, *Judicial Review in Equal Treatment Cases*, 137–42. Nilsson, 'Same, Same but Different', 129. Rory O'Connell, ed., 'European Principles of Interpretation and Application', in *Law, Democracy and the European Court of Human Rights* (Cambridge: Cambridge University Press, 2020), 72, https://doi.org/10.1017/9781139547246.006.

[437] ECtHR Fourth Section (19 February 2019), Case of Aleksić v. Slovenia, Application no. 57123/10.

[438] ECtHR (21 July 2016, 30 January 2017 Final) Case of Mamatas and Others v. Greece, Application nos. 63066/14, 64297/14 and 66106/14.

[439] ECtHR (16 March 2010) Case of Oršuš and Others v. Croatia, Application no. 15766/03.

[440] Original: "une certaine conception du 'vivre ensemble' » ECtHR Second Section (11 July 2017, Final 11 December 2017), Case of Belcacemi and Oussar v. Belgium, Application no. 37798/13

[441] ECtHR Fifth Section (23 March 2017, Final 23 June 2017), Case of Wolter and Sarfert v. Germany, Application nos. 59752/13 and 66277/13, § 69.

[442] Council of Europe, 'Guide on Article 14 of the European Convention on Human Rights and on Article 1 of Protocol No. 12 to the Convention: Prohibition of Discrimination.', 19. ECtHR Grand Chamber (22 March 2012) Case of Konstantin Markin v. Russia, Application no. 30078/06, § 127. ECtHR Fourth Section (16 November 2004, Final 16 February 2005), In reference to the Case of Ünal Tekeli v. Turkey, Application no. 29865/96, § 63.

[443] ECtHR Grand Chamber (22 March 2012) Case of Konstantin Markin v. Russia, Application no. 30078/06, § 143.

be disadvantageous to both women's careers and to men's family life.[444] The ECtHR's anti-stereotyping approach showcases a shift towards more a more substantive understanding of equality and non-discrimination, acknowledging the socio-relational harms imposed by differentiating measures.[445] It delegitimises measures that look to consolidate historically disadvantageous socio-relational (power) structures, which in effect limit people's life choices and perpetuate gender subordination.[446]

Proportionality

A differentiating measure must be proportionate; there must be a reasonable relationship of proportionality between the means employed and the aim to be realised.[447] When assessing the proportionality of a given measure, the ECtHR will evaluate to what extent less intrusive alternatives were available to achieve the aim pursued, and whether the aim was important enough to justify an interference.[448] Typically, such a proportionality assessment comprises several distinct components of inquiry.[449] First, a measure must have rational connection to the goal pursued. The measure must be capable or *suitable* to achieve the goal. Second, the rule must be *necessary*, meaning there is no less intrusive, but equally effective alternative.[450] Finally, the

---

[444] ECtHR Grand Chamber (22 March 2012) Case of Konstantin Markin v. Russia, Application no. 30078/06, § 141.

[445] The Court's anti-stereotyping approach, and its relevance to algorithmic decision-making will be further discussed as part of Chapter 4, Section 4.3.2 Anti-Stereotyping, as well as Chapter 6, Section 6.2 "Algortihmic Stereotyping and Deindividualisation"

[446] ECtHR Grand Chamber (22 March 2012) Case of Konstantin Markin v. Russia, Application no. 30078/06, § 141, read together with § 119. In the latter paragraph, the Court refers to the Third Party's Submissions.

[447] See for example: ECtHR Grand Chamber (19 December 2018) Case of Molla Sali v Greece, Application no. 20452/14, § 135; ECtHR First Section (24 October 2019, final 24 February 2020), Case of J.D. and A v. The United Kingdom, Application nos. 32949/17 and 34614/17, § 86. See also: Case *"Relating to Certain Aspects of the Laws on the Use of Languages in Education in Belgium" v. Belgium (Merits)"* App no 1474/62, 1677/62, 1769/63, 1994/63, 2126/64 (ECHR, 23 July 1968) (The Belgian Linguistic Case), § 10.

[448] Europäische Union, Europarat, and Europäischer Gerichtshof für Menschenrechte, eds., *Handbook on European Non-Discrimination Law*, 2018 edition, Handbook / FRA, European Union Agency for Fundamental Rights (Luxembourg: Publications Office of the European Union, 2018), 93.

[449] For an extensive, theoretical analysis of the proportionality test, see also: K. Moller, 'Proportionality: Challenging the Critics', *International Journal of Constitutional Law* 10, no. 3 (1 July 2012): 709–31, https://doi.org/10.1093/icon/mos024. For an extensive evaluation of the proportionality test in relation to the right to equality: Gerards, *Judicial Review in Equal Treatment Cases*, chap. 3.

[450] Moller, 711. For an example on the "fair balancing requirement" see: ECtHR Grand Chamber (12 February 2008), Case of Kafkaris v. Cyprus, Application no. 21906/04, § 161: "Article 14 does not prohibit distinctions in treatment which are founded on an objective assessment of essentially different factual circumstances and which, being based on

measure should not disproportionately burden right-holders. In this last step, the ECtHR enters into a fair balancing exercise between the interests of the decision-maker and the impact their decision has on people's interests.[451] Whereas these steps can be discerned within the ECtHR's case-law, the Court has nonetheless been criticised (here as well) for lacking clarity and consistency in the application of the proportionality review.[452]

The ECtHR's proportionality approach is often brought in connection to the margin of appreciation and "very weighty reasons" doctrines. Both affect the level of scrutiny typically applied by the Court vis-à-vis a differentiating measure, and its evaluation of a given measure's necessity, suitability, or balance.[453]

MARGIN OF APPRECIATION – Member States enjoy a certain margin of appreciation in "assessing whether and to what extent differences in otherwise similar situations justify a different treatment."[454] The margin of appreciation afforded to States depends upon the circumstances, the subject-matter and the background of the case.[455] If Member States operate within a narrow margin of appreciation, the ECtHR typically applies a higher level of scrutiny when analysing differential treatment. If the margin of appreciation is wide, the Court will be more lenient.[456]

---

the public interest, strike a fair balance between the protection of the interests of the community and respect for the rights and freedoms safeguarded by the Convention."

[451] Moller, 711. This is understood as proportionality in the strict sense.

[452] See for instance: Arnardóttir, *Equality and Non-Discrimination under the European Convention on Human Rights*, 15–16. Gerards, *Judicial Review in Equal Treatment Cases*, 145–49. Jeremy McBride, 'Proportionality and the European Convention on Human Rights.', in *The Principle of Proportionality in the Laws of Europe*, ed. Evelyn Ellis (Hart Publishing, 1999), 23–36, https://doi.org/10.5040/9781472561947. On the underlying difficulties applying a proportionality test onto violations of equality, see however: Guy Lurie, 'Proportionality and the Right to Equality', *German Law Journal* 21, no. 2 (February 2020): 174–96, https://doi.org/10.1017/glj.2020.8.

[453] Council of Europe, 'Guide on Article 14 of the European Convention on Human Rights and on Article 1 of Protocol No. 12 to the Convention: Prohibition of Discrimination.', 19. The margin of appreciation doctrine was first introduced in the Handyside case: ECtHR (7 December 1976) Case of Handyside v. the United Kingdom, Application no. 5493/72. Europäische Union, Europarat, and Europäischer Gerichtshof für Menschenrechte, *Handbook on European Non-Discrimination Law*, 93. ECtHR Grand Chamber (19 December 2018) Case of Molla Sali v Greece, Application no. 20452/14, § 143.

[454] See for example ECtHR Grand Chamber (19 December 2018) Case of Molla Sali v Greece, Application no. 20452/14; §136.

[455] ECtHR Grand Chamber (19 December 2018) Case of Molla Sali v Greece, Application no. 20452/14, § 143. Council of Europe, 'Guide on Article 14 of the European Convention on Human Rights and on Article 1 of Protocol No. 12 to the Convention: Prohibition of Discrimination.', 19.

[456] Koen Lemmens, 'The Margin of Appreciation in the ECtHR's Case Law: A European Version of the Levels of Scrutiny Doctrine?', *European Journal of Law Reform* 2, no. 3 (2018): 96. Lemmens describes the margin as the "device that allows

Various contextual factors can influence the margin appreciation, including the existence of a European consensus, or whether a measure was imposed in areas or domains in which decision-makers or judges are traditionally perceived as "better placed" to assess the need for a given measure. In addition, the ECtHR will also consider the nature or the seriousness of the interference, and more specifically, whether differential treatment touches upon the core values protected by the Convention.[457] Gerards has observed however that in recent years, the ECtHR makes use of "incrementalism" as an interpretative mechanic when it has to address "relatively new and potentially sensitive and divisive subject matter." In these cases, the Court will proceed in a cautious and circumscribed manner, gradually building towards more general international standards and principles.[458]

VERY WEIGHTY REASONS – In equality cases, certain differentiating measures may only be justified by very weighty or particularly serious reasons can be adduced. In these cases, the ECtHR typically concludes a violation of Article 14 ECHR took place. The very weighty reasons doctrine is grounds-based in the sense that its application is very closely tied to the differentiation trait or characteristic used as the basis for a decision-making procedure. In particular, the ECtHR references the very weighty reasons doctrine only in relation to a specific set of characteristics: race and ethnicity; nationality; (illegitimate) birth or origin; sexual orientation; mental or physical capabilities; gender and religion.[459] The reason as to why these traits are considered more

---

the Court to let States pursue public interests, albeit under European supervision." While this could be seen negatively, as promoting a State's interests over individual rights, Lemmens notes that it could also be seen positively as having "the potential to contribute to finding a social balance, thereby avoiding the interpretation and use of human rights in such a way that the common good or the social fabric is seriously endangered owing to an all too solipsistic interpretation of human rights."

[457] Gerards, 'Intensity of Judicial Review in Equal Treatment Cases'; Janneke Gerards, 'The Margin of Appreciation Doctrine, the Very Weighty Reasons Test and Grounds of Discrimination', SSRN Scholarly Paper (Rochester, NY: Social Science Research Network, 1 May 2017), https://doi.org/10.2139/ssrn.2875230.

[458] Gerards, 'Margin of Appreciation and Incrementalism in the Case Law of the European Court of Human Rights', 507–8. In this regard, the incremental approach might be more related to the function of the ECtHR as uncovering the ECHR's underlying "moral truth." See for instance: Letsas, 'Strasbourg's Interpretive Ethic'.

[459] The margin of appreciation and very weighty reasons doctrines are best seen as distinct. For example, in the Case of Jurčić (*infra)*, the ECtHR first considered that typically a wide margin of appreciation is given to States in social and economic policies. However, it then continued to observe that, even within these domains, policies which discriminate on the basis of sex, must advance very weighty reasons. ECtHR (4 May 2021) Case of Jurčić v. Croatia, Application no. 54711/15, § 64-65. On the relationship between both doctrines, see also: Gerards, 'The Margin of Appreciation Doctrine, the Very Weighty Reasons Test and Grounds of Discrimination'. The very weighty reasons doctrine can also be linked to the "legitimate aim" question. According to Gerards, in certain cases where the doctrine was invoked, the legitimate aim might have been legitimate or reasonable, but not sufficiently weighty. In this context, not only should the

suspicious than others differs.[460] For instance, the Court has observed that differential treatment based upon gender requires very weighty reasons as the promotion of gender equality is a major goal of the Council of Europe.[461] Differential treatment on the basis of sexual orientation requires particularly weighty reasons because this distinction "operates in this intimate and vulnerable sphere of an individual's private life."[462] While the very weighty reasons doctrine may be tied to specific characteristics, the grounds in question do not automatically trigger a heightened level of scrutiny. Instead, the underlying rationales advanced by the Court as to why the use of these traits are deserving of a stricter review are decisive.[463]

---

differentiating measure be proportionate, the aim must be proportionate too and " sufficiently weighty compared to the interests affected by the classification." Gerards, *Judicial Review in Equal Treatment Cases*, 142–43.

[460] The notion of "suspect" grounds is not without controversy however. See: Aaron Baker, 'Proportional, Not Strict, Scrutiny: Against a U.S. "Suspect Classifications" Model under Article 14 ECHR in the U.K.', *American Journal of Comparative Law* 56, no. 4 (1 September 2008): 847–94, https://doi.org/10.5131/ajcl.2007.0030.

[461] See for example in the case of gender equality: ECtHR Grand Chamber (22 March 2012) Case of Konstantin Markin v. Russia, Application no. 30078/06, § 127; ECtHR Chamber (22 February 1994), Case of Burghartz v. Switzerland, Application no. 16213/90, § 27.

[462] See for example: ECtHR Fourth Section (2 March 2010) Case of Kozak v. Poland, Application no. 13102/02, § 92.

[463] Gerards, 'The Margin of Appreciation Doctrine, the Very Weighty Reasons Test and Grounds of Discrimination', 9, 18–19. Regarding the "suspect" nature of differentiation grounds and the rationales for considering them as such, Gerards points out that suspectness "is a complex notion, which covers a variety of factors and rationales […] Importantly, most grounds of discrimination are not 'suspect' in all cases, which means their presence does not always lead to intensified review. It further appears that the rationales mentioned by the Court are different for the various grounds associated with the very weighty reasons test." It is also interesting to note that the grounds for which the ECtHR has required "very weighty reasons" are not always those grounds that have been explicitly listed within the Convention. For example, while sexual orientation is considered suspect, this ground was one that was added through case-law. Likewise, language, political opinion, national or social origin or property are explicit grounds, but their use may not necessarily require very weighty reasons. However, certain other grounds, that are related to historical disadvantage are not suspect, nor well protected. For instance, a person's socio-economic status, including poverty, while often a cause for discriminatory treatment, may still appear under protected in the ECtHR's case-law. See also: Sarah Ganty, 'Poverty as Misrecognition: What Role for Anti-Discrimination Law in Europe?', *Human Rights Law Review*, no. 00 (2021): 46. Laurens Lavrysen, 'Strengthening the Protection of Human Rights of Persons Living in Poverty under the ECHR', *Netherlands Quarterly of Human Rights* 33, no. 3 (1 September 2015): 293–325, https://doi.org/10.1177/016934411503300303. According to Arnardottir, in many cases concerning "non-suspect" grounds, the ECtHR will deal with these issues through other substantive Convention articles, such as the right to privacy. Oddný Mjöll Arnardóttir, 'Multidimensional Equality from within: Themes from the European Convention on Human Rights.', in *European Union Non-Discrimination Law: Comparative Perspectives on Multidimensional Equality Law.*, ed. Dagmar Schiek and Victoria Chege (London and New York: Routledge-Cavendish, 2009), 63. For instance, in the case of Lacatus v. Suisse, the ECtHR dealt with poverty and begging under Article 8 ECHR. ECtHR Third Section (19 January 2021, Final 19 April 2021), Case of Lacatus v. Switzerland, Application no. 14065/15.

4.3.2 Normative Guidance?

By examining the ECtHR's margin of appreciation and very weighty reasons doctrine, a set of contextual factors and rationales that determine the Court's level of review can be singled out.[464] These contextual factors in turn, allow for the identification of those inequalities, and acts of discrimination, that are (likely to be considered) harmful. They can therefore help to articulate the egalitarian objectives the ECtHR's wishes to protect, as well as the egalitarian harms it hopes to guard people against - within the scope of the Court's jurisdiction. This Section will take a closer look at these factors and rationales, and the goals they represent. As part of this analysis, this Section will also analyse their robustness in light of the risk of harm algorithms represent.[465]

The Margin of Appreciation doctrine varies in accordance to (among others) the existence of a European consensus, the context or domain in which a differentiating measure was taken (including whether a decision-maker was better placed), and the overall impact a measure has on the core values underlying the ECHR.[466] The very weighty reasons doctrine has only been invoked in relation to a limited set of suspect differentiation grounds. Still, Gerards believes that it is not the grounds, but the underlying factors and rationales that will be decisive for the ECtHR's

---

[464] The distinction between a Court's "levels of scrutiny" is used within American doctrine, and often compared to the European notion of margin of appreciation. Yet, differences between both approaches may be notable. See also: Lemmens, 'The Margin of Appreciation in the ECtHR's Case Law: A European Version of the Levels of Scrutiny Doctrine?' The actual and current functioning of the proportionality test, including the position of the margin of appreciation doctrine, within the ECtHR equality case-law have been subject to discussion. See for instance: Aaron A. Ostrovsky, 'What's So Funny about Peace, Love, and Understanding - How the Margin of Appreciation Doctrine Preserves Core Human Rights within Cultural Diversity and Legitimises International Human Rights Tribunals International Law', *Hanse Law Review* 1, no. 1 (2005): 47–64; Gerards, 'Margin of Appreciation and Incrementalism in the Case Law of the European Court of Human Rights'; Lemmens, 'The Margin of Appreciation in the ECtHR's Case Law: A European Version of the Levels of Scrutiny Doctrine?' Baker, 'Proportional, Not Strict, Scrutiny'; Nilsson, 'Same, Same but Different'.

[465] See Chapter 2 Section 2.2 3 Algorithmic Egalitarian Harms. In the following, I shall draw upon the analytical work of several authors who have tried to uncover the normative guiding points developed by the Court. Most notably: Gerards, 'Intensity of Judicial Review in Equal Treatment Cases'; Gerards, *Judicial Review in Equal Treatment Cases*; Gerards, 'The Margin of Appreciation Doctrine, the Very Weighty Reasons Test and Grounds of Discrimination'; Holzleithner, 'Mainstreaming Equality'; McColgan, 'Defining the Protected Characteristics'; Bamforth, 'Prohibited Grounds of Discrimination under EU Law and the European Convention on Human Rights'; Arnardóttir, 'The Differences That Make a Difference'; O'Connell, 'Cinderella Comes to the Ball'; Naudts, 'Criminal Profiling and Non-Discrimination: On Firm Grounds for the Digital Era?'

[466] See also: Gerards, 'The Margin of Appreciation Doctrine, the Very Weighty Reasons Test and Grounds of Discrimination'.

reasoning.[467] These rationales are (among others) the i) the irrelevance of the differentiating trait, ii) the immutability of the differentiating trait, ii) the stereotyping nature of the differentiating trait[468] or iv) the vulnerability of the groups captured by a specific ground.[469]

The ECtHR's actual conceptualisation of equality will be the result of a balancing exercise whereby the aforementioned factors can be used as indicators to gauge the justifiable nature of differentiation imposed onto people. Whereas some factors heighten the ECtHR's level of caution, others lower it; some gives Member States more discretion to differentiate, others do not. Not all criteria discussed below are "normative" in nature per se. Certain rationales clearly illustrate the ECtHR's shift towards a less formal and more substantive conceptualisation of equality (e.g., anti-stereotyping approach). Other factors are more closely tied to the objectives and goals pursued by the CoE and ECHR and constrain the ECtHR's capacity to intervene in national policies by imposing a particular conceptualisation of equality (e.g., better placed reasoning). In the end, egalitarian interests will be weighed against interests that are said to compete with equality, and the justification provided by, or discretion given to, the decision-maker, to override those interests. Bearing in mind the wider political context in which the ECtHR performs its review, the goal of the following Section is not to provide a determinate formula as to how this balancing exercise will be resolved. Instead, it is to uncover the egalitarian goals and objectives an ECtHR-inspired conceptualisation of equality embodies and what factors might heighten (or diminish) the role (that can be) afforded to equality.

European Consensus and Common Ground

In their interpretation of the fundamental right to equality and non-discrimination, the ECtHR must respond to any emerging consensus among Member States as to the standards of protection

[467] Gerards, 9; Arnardottir, 'The Differences That Make a Difference', 652–53.

[468] This includes situations where a stereotype lies at the foundation of the distinction or where the distinction in question (serves to) perpetuate(s) a stereotype.

[469] This uncoupling serves two purposes. First, these rationales typically reflect (substantive) ideals underlying the ECtHR's (normative) conceptualisation of equality. Second, by focusing on the conceptualisation of equality, it is easier to develop an evolutive understanding of equality within the ECtHR. More specifically, if we can assess to what extent those reasons are put forward as an argument to treat certain forms of differentiation with suspicion, they could (potentially) be applied to non-suspect grounds. In this regard, a differentiating measure is considered problematic, not because it is based upon a specific ground, but because it has come into conflict with a particular conceptualisation of equality.

those rights should achieve.[470] More specifically, the Court will "reassess 'in the light of present-day conditions' […] the appropriate interpretation and application of the Convention."[471] At first sight, the consensus approach allows the ECtHR to respond to challenges in a dynamic, evolutive and comparative manner, while safeguarding the principle of subsidiarity.[472] At the same time, the existence of consensus, or lack thereof, acts as a constraint towards the independent interpretation of equality by the ECtHR. Without consensus, the Court is more willing to accept justifications given to differential treatment. If there is consensus, the decision to move away from agreed upon standards is taken more seriously. [473]

The interpretation of human rights through consensus is not without criticism. Letsas has argued how consensus can lead to a situation where the "moralistic preferences of the majority" are imposed onto others. Those preference become "synonymous with 'public morals,'" which further legitimises the interests of the majority.[474] From an equality perspective, consensus can affect whether an individual or group can enjoy protection against egalitarian harms. According to Arnardottir, the consensus approach can "exclude or restrict pluralism as soon as individuals leave the cultural mainstream and find themselves located in the margins." [475] Biological sex equality may have entered the mainstream and is now an inherent component of ECHR's conceptualisation of equality, but issues of discrimination pertaining to gender, such as the rights of trans people or gender nonconforming persons, are relegated to the margins. As a result of a consensus approach, the disadvantage faced by a group must first become mainstream - or enjoy a degree of public understanding - if their claims are to enjoy protection.

---

[470] See for example: ECtHR Grand Chamber (28 May 2002), Case of Stafford v. the United Kingdom, Application no. 46295/99, § 68. ECtHR (16 November 2014), Case of Ünal Tekeli v. Turkey, Application no. 29865/96, § 54. Another famous example concerns the Handyside case where consensus was used in relation to 'public morals.' ECtHR (7 December 1976), Case of Handyside v. The United Kingdom, Application no. 5493/72.

[471] ECtHR Grand Chamber (28 May 2002), Case of Stafford v. the United Kingdom, Application no. 46295/99, § 69. "the Court proposes to reassess "in the light of present-day conditions" what is now the appropriate interpretation and application of the Convention." See also: Arnardottir, 'The Differences That Make a Difference', 650.

[472] ECtHR Grand Chamber (28 May 2002), Case of Stafford v. the United Kingdom, Application no. 46295/99, § 68. See also: Arnardottir, 650.

[473] Gerards, 'The Margin of Appreciation Doctrine, the Very Weighty Reasons Test and Grounds of Discrimination', 6–7. See for instance the Case of Dakir v. Belgium where the Court observed how there is still no consensus within Member States whether full-face veils can be banned in public places. ECtHR (11 July 2017), Case of Dakir v. Belgium, Application no. 4619/12, § 59.

[474] George Letsas, *A Theory of Interpretation of the European Convention on Human Rights* (Oxford: Oxford University Press, 2007), 121, https://doi.org/10.1093/acprof:oso/9780199203437.001.0001. These preferences are described as the "external preference that some people should not enjoy some liberty on the basis that their plan of life is inferior."

[475] Arnardóttir, 'Multidimensional Equality from within: Themes from the European Convention on Human Rights.', 59.

CONSENSUS AND ALGORITHMICALLY GUIDED DECISIONS - The defining role of consensus may therefore also have important consequences regarding the conceptualisation of equality vis-à-vis algorithmically guided decision-making processes. First, and related to Arnardottir's observation, these data-driven systems often have a disproportionate impact on people at the intersections and margins. Due to the social invisibility of their struggles, they often remain underrepresented within the data and design strategies underlying a decision-making process. As a consequence, the disadvantage they experience is often greater than other historically disadvantaged, but better recognised groups. Yet, a consensus conceptualisation of equality might be unable to capture and redress these specific harms.

Second, the algorithmic landscape is complex and fast-moving. Many of the challenges currently encountered are still novel. In such a rapidly transforming environment, consensus regarding the answers that must be given to algorithmic harms are unlikely to be found. This seems particularly the case for a legal system where differentiation grounds, rather than disadvantage, comprise the central conceptual component in defining equality. Unless attention shifts towards the actual disadvantage algorithmic decisions impose, it seems rather difficult to develop a sense of consensus regarding what algorithmic differentiation grounds should be evaluated and regulated, given their dynamic, intangible or seemingly random nature.[476] Given the novelty of this setting, perhaps an incremental approach, whereby the Court unearth the moral truth underlying the ECHR, by gradually setting standards and principles regarding the (un)justified use of novel differentiation grounds could prove valuable.[477]

The 'Better Placed' Argument, or the Domain of Application

The ECtHR tends to afford a wider margin of appreciation where national courts are deemed better placed or equipped to strike the correct balance between interests involved. There are two broad instances (of complex policy domains) where the ECtHR makes use of the better placed argument.[478] The first category concerns those areas characterised by a lack of consensus. The

[476] Of course, as highlighted in this Chapter's initial sections, new grounds have been added, so their analysis under Article 14 should not be excluded.

[477] Gerards and Zuiderveen Borgesius, 'Protected Grounds and the System of Non-Discrimination Law in the Context of Algorithmic Decision-Making and Artificial Intelligence', June 2022, 33.

[478] Gerards, 'The Margin of Appreciation Doctrine, the Very Weighty Reasons Test and Grounds of Discrimination', 7. G. Letsas, 'Two Concepts of the Margin of Appreciation', *Oxford Journal of Legal Studies* 26, no. 4 (1 January 2006): 722, https://doi.org/10.1093/ojls/gql030. Ganty, 'Poverty as Misrecognition: What Role for Anti-Discrimination Law in Europe?', 22. Oddný Mjöll Arnardóttir, 'Rethinking the Two Margins of Appreciation', *European Constitutional Law*

second instance concerns cases of great moral or political sensitivity. Furthermore, as an instrument safeguarding civil and political rights primarily, the Convention, in principle, does not directly deal with social rights, as well as material or economic distributions. The latter include areas that typically pertain to the domain of distributive justice, narrowly understood as those measures that relate to the distribution and allocation of resources within public policy: "Because of their direct knowledge of their society and its needs, the national authorities are in principle better placed than the international judge to appreciate what is in the public interest on social or economic grounds, and the Court will generally respect the legislature's policy choice unless it is 'manifestly without reasonable foundation'."[479]

For example, in the *Bah v. United Kingdom* judgment, the ECtHR found that in resource-hungry areas where there is insufficient supply to meet the demand, States can legitimately implement differentiating criteria to help guide the governance.[480] The ECtHR recognised that to be workable, welfare systems "may have to use broad categorisations to distinguish between different groups in need."[481] Despite this leniency, the ECtHR maintained that the differentiating criteria relied upon should not be arbitrary, nor should they constitute discrimination.[482]

Review 12, no. 1 (May 2016): 32, https://doi.org/10.1017/S1574019616000018. For more information on the 'better placed argument' and its relation to the margin of appreciation, see also the work of: Letsas, 'Two Concepts of the Margin of Appreciation'; Andrew Legg, *The Margin of Appreciation in International Human Rights Law: Deference and Proportionality*, Oxford Monographs in International Law (Oxford: Oxford University Press, 2012), https://doi.org/10.1093/acprof:oso/9780199650453.001.0001.

[479] ECtHR (4 May 2021) Case of Jurčić v. Croatia, Application no. 54711/15, § 64. In Jurčić, the matter at hand concerned discrimination within the allocation of pension schemes. See also: ECtHR Third Section (11 December 2018, Final 11 March 2019), Case of Belli and Arquier-Martinez v. Switzerland, Application no. 65550/13; ECtHR Fourth Section (27 September 2011, Final 27 December 2011), Case of Bah v. The United Kingdom, Application no. 56328/07, § 37; ECtHR Grand Chamber (6 July 2005), Stec and Others v. the United Kingdom, Application nos. 65731/01 and 65900/01, § 52

[480] In their dissenting opinion in the case of J.D. and A., judges Wojtyczeck and Pejchal note the following: "In our view, the judge and in particular the European judge, has to be extremely cautious in exercising his or her discretion in equality and discrimination cases, especially if their gist lies in an allegedly deficient allocation of financial resources." In other words, the Court must be careful as to not "leave the area of judicial enforcement of Convention rights and enter the field of policy-making." Joint Partially Dissenting Opinion of Judges Wojtyczek and Pejchal, ECtHR First Section (24 October 2019, final 24 February 2020), Case of J.D. and A. v. The United Kingdom, Application nos. 32949/17 and 34614/17, § 11.

[481] ECtHR Fourth Section (27 September 2011, Final 27 December 2011), Case of Bah v. The United Kingdom, Application no. 56328/07, § 44-47.

[482] In the case of Bah, the ECtHR considered that a government was justified in using people's immigrant status as a differentiation ground in the allocation of housing. The government argued that it was reasonable and justified to prioritise those who had a fixed and permanent right in order to ensure the fair allocation of a scarce resource. ECtHR

At the same time, the ECtHR acknowledges that even though the Convention "[sets] forth what are essentially civil and political rights, many of them have implications of a social or economic nature."[483] The interpretation of the Convention may indeed extend into the sphere of social and economic rights as there is no "water-tight division" to separate the civil and political sphere from the social and economic one.[484] Measures taken on social and economic grounds must therefore still satisfy the requirements of proportionality.[485] In this context, Mantouvalou has argued in favour of an integrated approach where the ECtHR can explore how "civil and political rights have inherent socio-economic components."[486] Policy questions that affect people's access to employment, their exclusion thereof and resulting poverty, ought to be addressed because they touch upon values underlying the ECHR, including equality, citizenship and social inclusion.[487]

ALGORITHMS - Algorithms are deployed in "politically sensitive domains", including the automation of a variety of social welfare decisions.[488] Within a grounds-based approach, algorithms pose a

---

Fourth Section (27 September 2011, Final 27 December 2011), Case of Bah v. The United Kingdom, Application no. 56328/07, § 45 -48.

[483] ECtHR Chamber (9 October 1979), Case of Airey v. Ireland, Application no. 6289/73, §26.

[484] ECtHR Chamber (9 October 1979), Case of Airey v. Ireland, Application no. 6289/73, §26. See for example also: Ingrid Leijten, 'Defining the Scope of Economic and Social Guarantees in the Case Law of the ECtHR', in *Shaping Rights in the ECHR: The Role of the European Court of Human Rights in Determining the Scope of Human Rights*, ed. Eva Brems and Janneke Gerards (Cambridge: Cambridge University Press, 2014), 109–36, https://doi.org/10.1017/CBO9781107337923.008.

[485] See also: ECtHR (4 May 2021) Case of Jurčić v. Croatia, Application no. 54711/15, § 64.

[486] Though Mantouvalou's focus lies primarily on the relationship between civil and political rights on the one hand, and social rights on the other hands. Virginia Mantouvalou, 'Work and Private Life: Sidabras and Dziautas v Lithuania', *European Law Review*, no. 30 (10 August 2005): 574.

[487] Virginia Mantouvalou, 584–85. In reference to: Hugh Collins, 'Discrimination, Equality and Social Inclusion', *The Modern Law Review* 66, no. 1 (2003): 16–43, https://doi.org/10.1111/1468-2230.6601002. Fredman also observed how in the D.H. Case, which concerned the education of Roma children (and as such, the touched upon the right to education, which is protected under Article 2 of Protocol 1 to the Convention for the Protection of Human Rights and Fundamental Freedoms (1952)) the ECtHR had focused primarily on the distributive, socio-economic disadvantage faced by the children. Whereas the Czech government justified their education policies as trying to accommodate the special needs of children, in effect the educational system resulted in a disproportionate allocation of Roma children to special schools, which often had an inferior curriculum. The distribution of a public prized good (namely education) was the main trigger for concern. Fredman, 'Emerging from the Shadows: Substantive Equality and Article 14 of the European Convention on Human Rights', 288. See ECtHR Grand Chamber (13 November 2007) Case of D.H. and Others v. The Czech Republic, Application no. 57325/00, § 193.

[488] See for instance the many examples in: 'Automating Society 2020' (AlgorithmWatch and BertelsmannSiftung, January 2021), https://automatingsociety.algorithmwatch.org/wp-content/uploads/2021/01/Automating_Society_Report_2020_-_Deutsche_Ausgabe.pdf.

twofold challenge in domains where governments are given great freedom to differentiate. First, algorithmic differentiation grounds often function as "broad categorisations." Moreover, in the case these grounds are novel and dynamic, and cannot be tied to a pre-established form of discrimination, their acceptance as a new form of discrimination is unlikely due to the aforementioned lack of consensus. Second, and unless they are based on spurious correlations, their data-driven nature makes it difficult to call them arbitrary.[489]

Both the consensus and better placed argument approach confront us with a similar problem. On the one hand, they run the risk that all non-representational algorithmic differentiation grounds are accepted by default unless they can be tied to a pre-established and socially non-accepted form of discrimination. While there is no consensus on their use, it is also difficult to call them arbitrary. On the other hand, if the judiciary would assess the use of non-representational algorithmic differentiation grounds through a more incremental approach it could give rise to criticisms of judicial activism or illegitimate policy making. If one is concerned with harms of exclusion in the face of important public policies, this is one instance where a clear proportionality test as a mechanism to regulate novel forms of differentiation would be desirable. It could be used to weigh the disadvantage experienced by an excluded group, against the needs of the government to differentiate.[490] In such a scenario, one could consider a specific algorithmic differentiation ground as "arbitrary" where, after having evaluated the suitability and necessity of the measure, the disadvantage for those involved clearly outweighs the benefits derived from differentiation.

Democracy, dignity, and autonomy

The more a differentiating measure touches upon the core of the ECHR or a Convention right, the less likely it will find justification.[491] Gerards identifies (at least) two key cores of the Convention:

---

[489] See also Chapter 5, Section 5.3.3 "Rationality, Arbitrariness and Reliability."

[490] This again points towards a danger commonly associated with algorithms, namely that the data-driven nature of these decisions is taken as reliable, objective, and rational too easily. In order to mitigate this, and while a large degree of discretion may still be given to the States, automated decisions must be presented in a sufficiently transparent and substantiated way. To have insight into the underlying logic of a system seems the only way through which for example the "broad categorisations" such a system uses can find justification. Leijten, 'Defining the Scope of Economic and Social Guarantees in the Case Law of the ECtHR', 124; Letsas, *A Theory of Interpretation of the European Convention on Human Rights*.

[491] Gerards, 'The Margin of Appreciation Doctrine, the Very Weighty Reasons Test and Grounds of Discrimination', 7. On the difficulty pinpointing the core of the Convention and Convention Rights, and the function this core holds in relation to the ECtHR's review, see also: Sébastien Van Drooghenbroeck and Cecilia Rizcallah, 'The ECHR and the Essence of

the "promotion and maintenance of the ideals and values of a democratic society" and the "protection of human dignity and personal autonomy."[492]

Certain forms of differentiation, such as racial discrimination, constitute a special affront to human dignity: the singling out of people on the basis of their race is degrading, particularly invidious, perilous and requiring special vigilance and vigorous reaction.[493] The ECtHR highlights how such attacks on human dignity undermine democratic values: any difference in treatment based exclusively on a person or group's ethnic origin, "can never be justified in a contemporary

---

Fundamental Rights: Searching for Sugar in Hot Milk?', *German Law Journal* 20, no. 6 (September 2019): 904–23, https://doi.org/10.1017/glj.2019.68.

[492] Gerards, 'The Margin of Appreciation Doctrine, the Very Weighty Reasons Test and Grounds of Discrimination', 7–8. In reference to *among others* the Case of Refah Partisi. Gerards, 'Intensity of Judicial Review in Equal Treatment Cases', 169–71. The nature of the interference may also be conceptualised in other ways. Arnardottir for example notes how direct discrimination and negative claims are typically subject to higher scrutiny than positive obligations and indirect discrimination. Arnardóttir, 'Rethinking the Two Margins of Appreciation', 57. See on the role of dignity also: C. McCrudden, 'Human Dignity and Judicial Interpretation of Human Rights', *European Journal of International Law* 19, no. 4 (1 September 2008): 655–724, https://doi.org/10.1093/ejil/chn043; Jean-Paul Costa, 'Human Dignity in the Jurisprudence of the European Court of Human Rights', in *Understanding Human Dignity*, ed. Christopher McCrudden (British Academy, 2013), 0, https://doi.org/10.5871/bacad/9780197265642.003.0022.

[493] ECtHR Second Section (13 December 2005, final 13 March 2006), Case of Timishev v. Russia, Application nos. 55762/00 and 55974/00, § 56. ECtHR Grand Chamber (10 May 2001) Case of Cyprus v. Turkey, Application no. 25781/94. *"With respect to an allegation of racial discrimination, that a special importance should be attached to discrimination based on race and that publicly to single out a group of persons for differential treatment on the basis of race might, in certain circumstances, constitute a special affront to human dignity. In the Commission's opinion, differential treatment of a group of persons on the basis of race might therefore be capable of constituting degrading treatment when differential treatment on some other ground would raise no such question."* The ECtHR referred to the Commission's report in the East African Asians case, the Court recalled that: ECtHR Grand Chamber (10 May 2001) Case of Cyprus v. Turkey, Application no. 25781/94, § 306; European Commission on Human Rights (14 December 1973) East African Asians v. the United Kingdom, Applications nos. 4403/70-4419/70, 4422/70, 4423/70, 4434/70, 443/70, 4476/70-4478/70, 4486/70, 4501/70 and 4526/70-4530/70, § 207. In the Nachova case, the Court called racial violence an "affront to human dignity." ECtHR Grand Chamber (6 July 2005) Case of Nachova and Others v. Bulgaria, Application nos. 43577/98 and 43579/98, § 145. This concerned the failure of state authorities to investigate possible racist motives against two Roma fugitives. Of course, several other convention rights could be seen as being directly tied to the protection of individual (equal) dignity and autonomy. The interference with a person's right to privacy and data protection, both protected under Article 8 ECHR, can be articulated as an interference with a person's dignity and autonomy. With regard to the relationship between equality and dignity, Réaume for example has proposed a dignity-based construction of equality rights. Denise Réaume, 'Dignity, Equality, and Comparison', in *Philosophical Foundations of Discrimination Law* (Oxford: Oxford University Press, 2013), https://doi.org/10.1093/acprof:oso/9780199664313.003.0002. . See for example Solanke, *Discrimination as Stigma*, 3. In reference to D Hellman and A Moreau (eds), Philosophical Foundations of Discrimination Law (Oxford , Oxford University Press, 2013).

democratic society built on the principles of pluralism and respect for different cultures."[494] The ECtHR is particularly critical where generalised assumptions and negative stereotypes linked to race and ethnicity are used as a basis for decision-making.[495] In this regard, Timmer has observed how stereotypes are rightfully condemned "because they restrict people to supposed group characteristics, thus impairing their dignity and personal autonomy, as well as denying them certain rights on this basis."[496]

The use of algorithms can come into conflict with notions of human dignity in many ways. First, inferential analytics could be deployed by (government) actors to discover, single out and target specific groups (and their members) within society. Data-driven technologies could be used to purposefully limit certain social groups in the exercise of their rights or make them subject to measures of surveillance or (targeted) physical or psychological violence. Motivated by homophobia for instance, a government could attempt monitor social media behaviour in order to identify members of the LGBTQI+ community.[497] Second, algorithmic decision-making systems have been found to replicate and reinforce stereotyped views that exist regarding certain (social) groups. For example, the Roma community has been subject to ethnic profiling throughout

[494] The ECtHR promotes diversity as an enrichment of democracy, rather than a threat thereto. ECtHR Second Section (13 December 2005, final 13 March 2006), Case of Timishev v. Russia, Application nos. 55762/00 and 55974/00, § 56-58. This case concerned a difference in treatment between persons of Chechen and non-Chechen ethnic origins. See also: ECtHR Fourth Section (16 April 2019) Case of Lingurar v. Romania, Application no, 48474/14.

[495] For example, in the Case of Lingurar, the ECtHR condemned a police raid based upon false stereotypes. The ECtHR observed: "The applicants' own behaviour was extrapolated from a stereotypical perception that the authorities had of the Roma community as a whole. The Court considers that the applicants were targeted because they were Roma and because the authorities perceived the Roma community as anti-social and criminal. This conclusion […], goes beyond a simple expression of concern about ethnic discrimination in Romania." In a similar fashion, the Court has condemned homophobic (and transphobic) differential treatment that comes into conflict with dignity. For example, in a case that concerned abusive police conduct during search of premises of an LGBT NGO motivated by homophobic and/or transphobic hatred, the ECtHR observed: "*treating violence and brutality with discriminatory intent, irrespective of whether they are perpetrated by State agents or private individuals, on an equal footing with cases that have no such overtones would be to turn a blind eye to the specific nature of acts that are particularly destructive of fundamental rights. A failure to make a distinction in the way situations that are essentially different are handled may constitute unjustified treatment irreconcilable with Article 14 of the Convention*." ECtHR Fifth Section (8 October 2020, Final 8 January 2021), Case of Aghdgomela Shivili and Japaridze v. Georgia.

[496] Timmer, 'Toward an Anti-Stereotyping Approach for the European Court of Human Rights', 709. See also Section 4.3.2 "Anti-Stereotyping"

[497] Whether these technologies can actually be accurate in their discovery of sensitive information is an altogether different question. See also: Fosch-Villaronga et al., 'A Little Bird Told Me Your Gender'.

Europe.[498] Where predictive policing systems are built upon police data imbued with historical prejudice, the system's outcomes will be extrapolated from the stereotypical perception that already existed.[499] Algorithmic systems thus impose an additional environment where certain nefarious forms of differential treatment find replication, and in so doing, impede upon people's dignity and autonomy. [500]

Moreover, algorithms may be deployed in ways that can come into conflict with the democratic values underlying the Convention, as well as each and every Convention right more generally.[501] Such violations may or may not involve a violation of equality and non-discrimination: opaque

---

[498] Europäische Union, ed., *Understanding and Preventing Discriminatory Ethnic Profiling: A Guide ; towards More Effective Policing* (Luxembourg: Office for Official Publ. of the Europ. Union, 2010); Sara Benedi Lahuerta, 'Ethnic Discrimination, Discrimination by Association and the Roma Community: CHEZ', *Common Market Law Review* 53, no. 3 (2016): 797–818; 'Ethnic Profiling: A Persisting Practice in Europe', Commissioner for Human Rights, accessed 21 September 2022, https://www.coe.int/en/web/commissioner/blog/-/asset_publisher/xZ32OPEoxOkq/content/ethnic-profiling-a-persisting-practice-in-europe.

[499] See also footnote 519, and the reference made to the Lingurar Case.

[500] More generally however, algorithmically generated knowledge can often can be reformulated as a generalised statement or assumption concerning the algorithmic group (and its members). They too restrict group members to their supposed characteristics, and therefore fail to recognise a person's (or group's) dignity and autonomy. When they apply to historically disadvantaged groups however, these generalised assumptions might be even more difficult to challenge. In Chapter 6 and 7, it will also be explained in more detail how algorithms in their failure to guarantee the moral, socio-relational or distributional equality through their functioning could violate the dignity and autonomy of decision-making subjects. In particular, algorithms have the capacity to undermine people in the exercise of their right to self-determination and self-development. This analysis will also consider the use of algorithmic generalisations.

[501] Rather than looking at all civil and political rights in isolation, and investigate to what extent they are democratically necessary, one could argue that, more generally speaking, all fundamental rights and freedoms protected under the ECHR could be considered as pertaining to the core ideals and values a democratic society. In this regard, the right to equality and the prohibition of discrimination protected under Article 14 ECHR is the distributive mechanism that ensures the equal access and enjoyment of the Convention's fundamental freedoms. Article 14 ensures the equal distribution of the "prized goods" that are the fundamental rights. McCrudden and Kountouros, 'Human Rights and European Equality Law'; McCrudden and Prechal, 'The Concepts of Equality and Non-Discrimination in Europe: A Practical Approach'. In the Refah Partisi Chamber judgement (before it was referred to the Grand Chamber), the ECtHR also clarified the relationship between the rule of law, and equality before the law: "The rule of law means that all human beings are equal before the law, in their rights as in their duties. However, legislation must take account of differences, provided that distinctions between people and situations have an objective and reasonable justification, pursue a legitimate aim and are proportionate and consistent with the principles normally upheld by democratic societies. But the rule of law cannot be said to govern a secular society when groups of persons are discriminated against solely on the ground that they are of a different sex or have different political or religious beliefs. Nor is the rule of law upheld where entirely different legal systems are created for such groups." ECtHR Third Section (31 July 2001) Case of Refah Partisi (The Welfare Party) and Others, Applications nos. 41340/98, 41342/98, 41343/98 and 41344/98, §43.

predictive policing systems interfere with a person's right to fair trial; biased automated voting systems can upset the democratic electoral process, etc. Given that differentiation is inherent to algorithmic decision-making processes, there is an increased risk that automated processes unjustifiably violate the equal enjoyment of the rights and freedoms protected under the ECHR. For decision-makers, the Convention assumes a signalling function: wherever and whenever automated practices are deployed in ways that can affect the enjoyment of a given fundamental freedom or right, any type of differential treatment should be met with greater caution. If the system relies upon differentiation, they should assess to what extent the decision-making process can result in the exclusion of a particular person or group regarding the equal enjoyment or access to the protection of a Convention right. As soon as someone is precluded, for instance because the system appears biased, they should verify whether this exclusion can be justified. For example, if a State-run media platform allows political parties to run targeted campaigns, they should be able to assess how and under what conditions, these parties target their audience, and to what extent non-targeted people can still attain sufficient information.

Immutability

The notion of immutability typically pertains to characteristics that cannot be chosen and changed, such as a person's ethnicity or biological sex. However, it can also refer to traits that are so closely related to a person's identity or sense of self that they cannot be considered the subject of choice. Instead, immutable traits are an integral component of one's personality.[502] A person's religious beliefs for instance can viewed as a core aspect of one's identity.[503]

Immutability has long played a key role in grounding non-discrimination laws. Following the immutability rationale, the application of the anti-discrimination principle is politically legitimate when it protects people against mistreatment for reasons over which they have little or no control

[502] See also: Solanke, *Discrimination as Stigma*, 54–59. SA Marcosson , ' Constructive Immutability ' (2001) 3 (2) University of Pennsylvania Journal of Constitutional Law 673 .

[503] In this regard, Gerards and Borgesius point towards the distinction between the forum externum (or the expression of religion) on the one hand, which could be more reasonably considered as chosen, and the forum internum (the holding of a religion), for which the notion "choice" is more difficult to uphold. Gerards and Zuiderveen Borgesius, 'Protected Grounds and the System of Non-Discrimination Law in the Context of Algorithmic Decision-Making and Artificial Intelligence', June 2022, 51.

or are impossible or exceedingly difficult to change.[504] Conversely, people can be differentiated on the basis of traits and characteristics they do have control over.[505]

Even though the ECtHR appears to have abandoned immutability as a definitional criterion for discrimination, the notion has still found use when determining the Court's level of review.[506] In the aforementioned case of *Bah v. the United Kingdom*, the Court found that a person's immigration status was not an immutable trait.[507] This observation was significant during the review phase: "Immigration status is not an inherent or immutable personal characteristic such as sex or race, but is subject to an element of choice […] Given the element of choice involved in immigration status, therefore, while differential treatment based on this ground must still be objectively and reasonably justifiable, the justification required will not be as weighty as in the case of a distinction based, for example, on nationality."[508]

---

[504] Solanke, *Discrimination as Stigma*, 55. This "historical association" between immutable traits and non-discrimination law has also been discerned within the EU system. See for example: Mark Bell and Lisa Waddington, 'Similar Yet Different: The Work-Life Balance Directive and the Expanding Frontiers of EU Non-Discrimination Law.', *Common Market Law Review* 58, no. 5 (2021): 1410.

[505] See also: Arnardóttir, 'Multidimensional Equality from within: Themes from the European Convention on Human Rights.', 56.

[506] See Section 4.2 "The Range of Protection". The ECtHR argued that: "different treatment […] was not based on any personal choice in so far as this choice should be respected as elements of someone's personality, such as religion, political opinion, sexual orientation and gender identity, or on grounds of personal features in respect of which no choice at all can be made, such as sex, race, disability and age." ECtHR (22 November 2012; Final 29 April 2013) Case of Cadek and Others v. The Czech Republic, Applications nos. 31933/08, 60084/08, 6185/09, 46696/09, 52792/09, 53518/09, 10185/10, 42151/10, 3167/11 and 20939/11, § 94. In Peterka, the Court observed that a person's contract status is subject to change and therefore should not be afforded the same protection against differentiation on the basis of a ground that either represents a personal choice to in shaping one's identity or personality, or of a ground over which a person has no choice at all. ECtHR (4 May 2010) Josef Peterka v. Czech Republic, Application no. 21990/08, p. 6.

[507] ECtHR Fourth Section (27 September 2011, Final 27 December 2011), Case of Bah v. The United Kingdom, Application no. 56328/07, § 45.

[508] ECtHR Fourth Section (27 September 2011, Final 27 December 2011), Case of Bah v. The United Kingdom, Application no. 56328/07, § 47. In this regard, the ECtHR observed that for public welfare service, especially those that are "resource-hungry", broad categorisations and criteria, such as migration status, serve a legitimate purpose. The Court concluded that: "that the differential treatment to which the applicant was subjected was reasonably and objectively justified by the need to allocate, as fairly as possible, the scarce stock of social housing available in the United Kingdom and the legitimacy, in so allocating, of having regard to the immigration status of those who are in need of housing."

The immutability rationale has been linked to the liberal idea that people are autonomous and rational actors with full control over their choices.[509] To take a person's alleged control over choice as a key element in determining the justified nature of the inequalities they incur, can be criticised however. Choices are always made under certain conditions and people do not always have control over the latter. These conditions too should be taken into consideration when determining whether or not a given decisional outcome was justified.[510] It is unreasonable to hold people (fully) responsible for their inequalities on the basis of choice alone.[511] Moreover, it is difficult to actually define the immutability of, or choice exercised in relation to, a given characteristic. To what extent is a person's immigration status subject to choice, rather than conditioned by the particular circumstances a person finds themselves in? The immutability rationale may very well protect against certain invidious forms of discrimination, such as sexism or racism, but given its shaky foundations, the rationale might exclude other vulnerable groups and stigmatised identities from protection.[512] Furthermore, to focus on the immutability of a characteristic is to focus on the nature of the differentiating trait, rather than the disadvantage linked to it.[513] Put differently, the rationale is more closely associated to formal approaches towards equality, rather than goal-oriented or transformative ones. Finally, in many instances immutable and non-choice traits, such as a person's natural talents and abilities, are considered legitimate grounds for decision-making.[514] Put simply, the immutability rationale fails to adequately substantiate equality and non-discrimination laws.

IMMUTABILITY AND ALGORITHMS - The application of immutability, choice, and control as determinant factors for assessing the justified nature of algorithmic inequalities is equally

---

[509] Alexandra Timmer, *Strengthening the Equality Analysis of the European Court of Human Rights: The Potential of the Concepts of Stereotyping and Vulnerability.*, 2014, 235.

[510] Nicholas Barry, 'Defending Luck Egalitarianism', *Journal of Applied Philosophy* 23, no. 1 (2006): 89–107, https://doi.org/10.1111/j.1468-5930.2006.00322.x.

[511] See also: Peter Vallentyne, 'Brute Luck, Option Luck, and Equalityof Initial Opportunities', *Ethics* 112, no. 3 (April 2002): 529–57, https://doi.org/10.1086/339275; Shlomi Segall, *Why Inequality Matters: Luck Egalitarianism, Its Meaning and Value* (Cambridge: Cambridge University Press, 2016), https://doi.org/10.1017/CBO9781316416969; Barry, 'Defending Luck Egalitarianism'; Kristin Voigt, 'The Harshness Objection: Is Luck Egalitarianism Too Harsh on the Victims of Option Luck?', *Ethical Theory and Moral Practice* 10, no. 4 (1 August 2007): 389–407, https://doi.org/10.1007/s10677-006-9060-4.

[512] Solanke, *Discrimination as Stigma*, 61.

[513] See e.g.: Timmer and Peroni, 'Bah v UK: On Immigration, Discrimination and Worrisome Reasoning'. In reference to J. M. Balkin, 'The Constitution of Status', *The Yale Law Journal* 106, no. 8 (June 1997): 2313, https://doi.org/10.2307/797222.

[514] See also: Gerards and Zuiderveen Borgesius, 'Protected Grounds and the System of Non-Discrimination Law in the Context of Algorithmic Decision-Making and Artificial Intelligence', June 2022, 50.

delicate.[515] When does an algorithmic differentiation ground reflect choice? Should it be prohibited to categorise people on the basis of data that do not mirror their choice? The COMPAS algorithm gave defendants a recidivism score on the basis of the answers they provided to a questionnaire.[516] One of those questions gauged whether the friends of the defendant used drugs. Even though a person has no (immediate) control over their friends' activities, it could be argued that they do have some degree of control over their choice of friends. Would this element of control be sufficient to render the use of said characteristic justified?

Consider a similar complex profile deemed relevant within a government's surveillance policy: "people with religion A who travel to the airport at night via Uber should be made subject to heightened surveillance." Does this combined differentiation ground represent an immutable characteristic? Religion may be an immutable trait, but all the other attributes within the profile at large (time of travel and mode of transportation) constitute autonomous choices over which people do exercise control. Suppose a government web portal helps people re-enter the job market. For the allocation of aid resources, the government differentiates between individuals on the basis of their viewing behaviour during the navigation of the online web application. They have discovered that persons who spend more time browsing the various job opportunities have a higher chance of success. The government uses this information to their advantage. Resources are more efficiently spent when they benefit those who are more likely to generate a future economic return. To what extent does browsing behaviour represent a person's choice or control?

Many algorithmic decisions take into account attributes and characteristics over which people have some degree of control.[517] Do all these instances of algorithmic differentiation correspond to our initial understanding of choice and non-choice as understood under the immutability rationale? The argument has also been made that due to the opacity and complexity of algorithmic systems, people have no meaningful choice or control over the inequalities they incur as a result of their deployment. Unless adequate mechanisms have been put in place that either enable decision-subjects to reasonably foresee the consequences and impact of their choices, actions and behaviours, or have given them enhanced control over the criteria used within a decision-making system, algorithmic differentiation grounds should be considered immutable, and their use

[515] Naudts, 'Fair or Unfair Differentiation? Luck Egalitarianism as a Lens for Evaluating Algorithmic Decision-Making.'

[516] See also Chapter 1, Section 1.2.2 "The Complexity of Decision-Making Grounds" Northpointe, 'COMPAS Risk & Need Assessment System'; Mattu et al., 'How We Analyzed the COMPAS Recidivism Algorithm'; Store, 'COMPAS Recidivism Risk Score Data and Analysis'.

[517] Of course, there might also be actions, behaviours or preferences over which decision-subjects clearly hold no control, such as subconscious actions and behaviours covertly monitored by algorithmic systems.

unjustified.[518] In conclusion, within the algorithmic environment the boundary between what is chosen and not-chosen has become increasingly blurred.[519]

Irrelevance

Under the irrelevancy rationale, differential treatment will be unjustified when it was based upon a characteristic that should have been irrelevant for the decision at hand, or where it considered that the characteristic should not affect a person's life opportunities.[520] The irrelevancy (or arbitrary) rationale can be linked to the formal approach towards equality and the performance of the traditional comparability test. As above-mentioned, under a formal approach, discrimination occurs when there is a difference in treatment of persons in analogous or relevantly similar situations. In order to assess whether people are in analogous situations, the Court will engage into a comparative review and identify a specific feature of relevance.[521] In particular, the Court will need to identify whether the differentiation ground relied upon by a decision-maker was relevant for the decision at hand. If no relevant difference between those involved can be identified, the measure was imposed onto persons in a relevantly comparable situation. In her analysis of the case-law, Gerards observed how in (many, though not all) cases where the basis for differentiation was a suspect characteristic, the Court often did not engage into an intensive comparative review. Suspect characteristics are typically perceived as bearing little social relevance for decision-making. In these instances, the ECtHR is more likely to assume that persons affected by a differentiating measure based on these traits are in a comparable situation.[522]

---

[518] In this context, Wachter has introduced the notion of "artificial immutability": "Insofar as immutability concerns personal control, arguments concerning the legal status of immutability apply as much to traditional immutable characteristics as they do to AI-defined artificially immutable characteristics. 'Artificial immutability' […] only posits that opacity, vagueness, instability, and a lack of personal control and social meaning render a characteristic *de facto immutable."* This begs the question whether, without mechanisms of control, the usage of such algorithmically generated grounds can be morally or legally justifiable. Naudts, 'Fair or Unfair Differentiation? Luck Egalitarianism as a Lens for Evaluating Algorithmic Decision-Making.' See also: Wachter, 'The Theory of Artificial Immutability', 48.

[519] In this regard, one can wonder whether one should also link a level of reasonable expectations in order to speak of choice and control and responsibility?

[520] Gerards, 'Intensity of Judicial Review in Equal Treatment Cases', 165–66. See also Chapter 3, Section 3.1.2. "Formal versus Substantive Equality." Khaitan, *A Theory of Discrimination Law*, 56.

[521] Cartabia, 'The European Court of Human Rights', 812.

[522] Gerards, *Judicial Review in Equal Treatment Cases*, 131–32.

IRRELEVANCE AND ALGORITHMS - In most cases it will be rather difficult to claim the alleged irrelevance of an algorithmic differentiation ground.[523] Indeed, the irrelevance rationale seems completely at odds with the entire purpose and function for which modern knowledge discovery tools are used. The relevance of a given differentiation ground lies entirely in the eye of the beholder. Because they have access to individualised information, decision-subjects may argue that a given differentiation ground was irrelevant. The decision-maker, having access to vast amounts of data, can easily claim the opposite.[524] Furthermore, like immutability, to make a sweeping claim regarding the irrelevancy of a differentiating trait, it to make a claim regarding the nature of that ground, rather than the disadvantage (socio-relational or distributive) linked to its use.[525] In other words, taking an irrelevance rationale as a point of departure, it becomes more difficult to argue that the inequalities imposed by algorithms, as well as the harms associated thereto, are unjustified.[526]

The Anti-stereotyping Approach

Rationales based upon the immutable, irrelevance, (or arbitrary,) nature of differentiation grounds represent formal equality. At the same time, the ECtHR has moved towards a more substantive and transformative, socio-relational, and distributional conceptualisation of equality. This transition is exemplified in the Court's development of an anti-stereotyping approach, and the use of the notion "vulnerability" to highlight the precarious social position certain individuals and groups find themselves in. Both rationales represent the ECtHR's willingness to engage with (historical) prejudice and disadvantage as contextual factors during the evaluation of differential treatment. Rather than focusing on the *nature* of the traits used, in cases where the ECtHR invokes an anti-stereotyping or vulnerability rationale, attention shifts towards the social causes that underlie moral, socio-relational (and even distributional) inequalities.[527]

---

[523] See above, Section 4.3.2. "The Better placed Argument"

[524] See also Chapter 6, Section 6.2 "Algorithmic Stereotyping and Generalisations"

[525] See above, Chapter 3, Section "Formal and Substantive Equality." Fredman, 'Providing Equality', 164.

[526] Wachter, 'The Theory of Artificial Immutability', 20–21.

[527] For a more thorough analysis of the anti-stereotyping approach of the ECtHR, reference can be made to the work of Timmer. Timmer, *Strengthening the Equality Analysis of the European Court of Human Rights: The Potential of the Concepts of Stereotyping and Vulnerability.*; Timmer, 'Toward an Anti-Stereotyping Approach for the European Court of Human Rights'; Timmer, 'Judging Stereotypes'; Alexandra Timmer, 'Gender Stereotyping in the Case Law of the EU Court of Justice', no. 1 (2016): 10; Eva Brems and Alexandra Timmer, eds., *Stereotypes and Human Rights Law* (Intersentia, 2016), https://doi.org/10.1017/9781780685458. For an exploration of the concept stereotype in relation to algorithms, see also Chapter 3 of this dissertation.

The above-mentioned case of *Konstantin Markin v. Russia* is considered a landmark anti-stereotyping case.[528] The ECtHR observed how laws founded on generalised assumptions or stereotypes can perpetuate social inequality and disadvantage and consolidate traditional relationships of power. As a reminder, this case concerned a Russian law which denied parental leave to male military personnel. [529] Not only should the law avoid the replication of (negative) stereotypes, those stereotypes - even when they are reflective of certain traditions, such as the traditional distribution of gender roles within society - cannot constitute a legitimate aim.[530] In the case of *Carvalho Pinto de Sousa Morais v. Portugal*, the ECtHR rightfully criticised a national judge's reliance on (gender) stereotypes in legal reasoning.[531] The Court explicitly acknowledged how stereotypes can undermine personal autonomy and personal dignity. In this particular age, the stereotyped assumption that sexuality holds lesser importance for women denies the "physical and psychological relevance for the self-fulfilment of women."[532] Finally, in the case of *Lingurar*, the ECtHR found that stereotypes cannot be relied upon to inform decisions taken by actors who represent the State, such as law enforcement agencies.[533] Ethnic profiling strategies discriminate through an extrapolation onto individuals of a stereotypical perception whereby criminality (and anti-social behaviour) is linked to a group's ethnicity.[534] The anti-stereotyping

---

[528] See also *supra* Section 4.3.1 "Legitimate Aim." ECtHR Grand Chamber (22 March 2012) Case of Konstantin Markin v. Russia, Application no. 30078/06

[529] The law was based upon gender stereotypes: women were seen as the primary child-carers and men as breadwinners. These stereotyped views were not only considered damaging towards the career opportunities of women, but they would also negate men the enjoyment of their family life. The Court was therefore unconvinced by the argument that a difference in treatment should have been considered a measure of affirmative action: "different treatment of servicemen and servicewomen as regards entitlement to parental leave [was] clearly not intended to correct the disadvantaged position of women in society or "factual inequalities" between men and women." ECtHR Grand Chamber (22 March 2012) Case of Konstantin Markin v. Russia, Application no. 30078/06, § 141.

[530] The Court was also unconvinced by the government's argument that (Russian) society would be unwilling to accept equality between men and women serving in the army. ECtHR Grand Chamber (22 March 2012) Case of Konstantin Markin v. Russia, Application no. 30078/06, § 142-143

[531] In this particular case, the stereotype concerned the wrongful assumption that sexuality was not as important for a fifty-year-old woman and mother of two. According to the Court, such reflected a traditional idea regarding female sexuality, and the supposed role age plays therein. That is sexuality is not as important for older women, and that sexuality is mainly linked to child-bearing purposes However, the ECtHR failed to explicitly name the ageist stereotype that a person's sexuality can be linked to their age. Peroni, 'Age and Gender Discrimination: Laudable Anti-Stereotyping Reasoning in Carvalho Pinto V. Portugal'. ECtHR (25 July 2017, Final 25 October 2017) Case of Carvalho Pinto de Sousa Morais v. Portugal, Application no. 17484/15, § 52.

[532] ECtHR (25 July 2017, Final 25 October 2017) Case of Carvalho Pinto de Sousa Morais v. Portugal, Application no. 17484/15, § 52.

[533] ECtHR Fourth Section (16 April 2019) Case of Lingurar v. Romania, Application no, 48474/14, § 76.

[534] ECtHR Fourth Section (16 April 2019) Case of Lingurar v. Romania, Application no, 48474/14, § 76.

approach clearly mandates the avoidance of generalised assumptions and negative stereotypes throughout the entire decision-making chain. Stereotypes should not justify the adoption of a law, nor should they impact the judgments of decision-makers, whether they are the law makers themselves, the judiciary, or executive forces, such as law enforcement agencies. In addition, decision-makers should avoid that their decisions reinforce stereotyped beliefs.

The anti-stereotyping approach can be justified in reference to the nefarious effects stereotypes have on people's social relationships, and their ability to maintain and reinforce structures of oppression and domination. Stereotypes emerge from, and help to maintain, the negative perceptions and attitudes that exist towards certain groups within society. The ECtHR considers as particularly invidious the use of stereotypes that are either untrue, based upon prejudice, or both.[535] In addition, to treat people on the basis of generalised assumptions is to violate their human dignity and personal autonomy.[536] Stereotypes attach a label to a person or group, which they must actively combat (and overcome) in their social navigations and relations. While this in itself already constrains people in their individuality and exercise of choice, due to the negative perceptions attached to their person, those stereotyped often find themselves excluded from being able to meaningfully participate within society and from enjoying certain socially prized or justice-relevant good, which in turn diminishes their capacity to exercising autonomous choice is further diminished.

The ECtHR however has also been criticised for not always recognising stereotyped reasoning, and even being complicit in reinforcing certain stereotypes, such as victim-blaming in cases of gender and domestic violence.[537] At the same time, it should be acknowledged that it is far from

---

[535] See also: Timmer, 'Judging Stereotypes', 247–50. As stereotypes often concern prejudicial overgeneralisations, they tend to be untruthful to some extent. For instance, in Kiyutin v. Russia a man was refused a residence permit on account of him being HIV positive and in order to prevent the further transmission of HIV. This, concluded the Court, was founded upon the assumption that HIV-positive persons engage in more unsafe behaviour. The latter was a generalisation however, not founded in any factual basis. ECtHR First Section (10 March 2011, Final 15 September 2011), Case of Kiyutin v. Russia, Application no. 2700/10, § 68.

[536] On dignity, see also Section 4.3.2 Democracy, Dignity and Autonomy. In the recent Jurčić v Croatia case, the Court too was worried about the domestic authorities' gender stereotyping "overtones." This direct discrimination case, which concerned the refusal of an employment-related benefit to pregnant women based on their pregnancy, the Court criticized the assumption that pregnant women should not seek employment while pregnant. ECHR (4 May 2021) Case of Jurčić v Croatia, Application no. 54711/15, § 83.

[537] Timmer, *Strengthening the Equality Analysis of the European Court of Human Rights: The Potential of the Concepts of Stereotyping and Vulnerability.*; Ganty, 'Poverty as Misrecognition: What Role for Anti-Discrimination Law in Europe?' Weinberger, 'Kurt v. Austria: A Missed Chance to Tackle Intersectional Discrimination and Gender-Based Stereotyping

straightforward to recognise stereotypes, and the various ways in which stereotypes can alter policy making. In addition, there is also a risk that the Court does not see stereotypes as a problem of equality, but instead analyses the issue a under different substantive right.[538] To consider what course of action would be appropriate, can therefore be difficult. For example, how to deal with the use of negative stereotypes that nonetheless result in more favourable treatment of an otherwise disadvantaged group?[539] This was the case of *Khamtokhu and Aksenchick v. Russia* where a male applicant argued a violation of Article 14 ECHR because men could face life imprisonment whereas women could not.[540] The government had argued that the differential treatment between men and women was justified not only due to the physiological differences between both sexes, but due to the special role women hold with regard to motherhood and reproduction.[541] Despite the government's reasoning being based upon generalised assumptions, the ECtHR argued there had been no violation of Article 14 ECHR. In her concurring opinion, Judge Turković pointed out that one could nonetheless "[…] criticise the majority for carrying out a

---

in Domestic Violence Cases'. Also within the ECtHR case-law this criticism has sometimes been noticeable. For example, in his concurring opinion in the Biao v. Denmark Case, Judge Pinto De Albuquerque, more explicitly refers to the negative impact of wrongful stereotypes: *"Ultimately, the stereotype that resident foreigners and Danish nationals of foreign origin are helpless young people, who are either forced to marry persons from their country of origin or tend to engage in an odd, "widespread" marriage pattern of a kind of cultural in-breeding, and later on build "unhappy" families, have "marital problems" and do not integrate well into society, is not confirmed by any objective evidence. In sum, the contested policy […] is based on a confused amalgam of misguided, biased assumptions which portray a surreal image of resident foreigners and Danish nationals of foreign origin living in Denmark, and more specifically – and most disturbingly – of those coming from "non-Western countries", in contrast with an idealised image of ever-faithful Danes, born in Denmark, who live outside the country. To put it bluntly, the Government's case is not weak on the facts, it is simply not made out."* Concurring opinion of Judge Pinto De Albuquerque in ECtHR Grand Chamber (24 May 2016) Case of Biao v. Denmark, Application no. 38590/10. The ECtHR did however refer to their Konstantin Markin conclusions that biased assumptions and prevailing social prejudice do not provide sufficient justification to impose a difference in treatment on grounds of sex. I*bid, § 126.* Weinberger.

[538] Timmer, *Strengthening the Equality Analysis of the European Court of Human Rights: The Potential of the Concepts of Stereotyping and Vulnerability.*, 223.

[539] For an illustration on the difficulty of naming and contextualising stereotypes, reference can be made to the concurring and dissenting opinion: ECtHR Grand Chamber (24 January 2017) Case of Khamtokhu and Aksenchick v. Russia, Application nos. 60367/08 and 961/11. See also: Corina Heri, 'Between a Rock and a Hard Place: The Court's Difficult Choice in Khamtokhu and Aksenchik v. Russia', *Strasbourg Observers* (blog), 17 March 2017, https://strasbourg.weichie.dev/2017/03/17/between-a-rock-and-a-hard-place-the-courts-difficult-choice-in-khamtokhu-and-aksenchik-v-russia/.

[540] ECtHR Grand Chamber (24 January 2017) Case of Khamtokhu and Aksenchick v. Russia, Application nos. 60367/08 and 961/11. One consequence could have been that if the Court had indeed found a violation of Article 14 ECHR, the reintroduction of life imprisonment for all could be proposed as a solution. See also: Heri.

[541] ECtHR Grand Chamber (24 January 2017) Case of Khamtokhu and Aksenchick v. Russia, Application nos. 60367/08 and 961/11, § 47.

scant analysis of equality and gender issues and for avoiding a discussion of possible stereotypes and their implications […] In my opinion, the Court should not refrain from naming different forms of stereotyping and should always assess their invidiousness. It is impossible to change reality without naming it."[542]

THE ANTI-STEREOTYPING APPROACH AND ALGORITHMS – The anti-stereotyping perspective encourages decision-makers to evaluate to what extent the systems they employ, reflect generalised ways of thinking about people, whether during creation or upon deployment.[543] First, decision-makers should avoid that algorithmic differentiation grounds and algorithmic decision-making processes reflect particular stereotypical views. For example, an algorithm that allocates a higher risk-score to persons with an HIV-positive status should not do so building upon the assumption that people with HIV engage in more unsafe behaviour. Second, even in the case an algorithm does not depart from prejudicial or untruthful assumptions, but from verifiably accurate data-driven research, decision-makers should still make sure that the deployment of algorithms does not introduce or perpetuate generalised assumptions concerning certain (social) groups. In the HIV example, and if HIV would indicate a risk, a decision-maker should be transparent and provide the scientific reasoning as to why this would be the case, and why it would be justified to make use of this status. Finally, the ECtHR's conceptualisation of equality is focused on negative stereotypes attached to historically disadvantaged groups. If the approach should be understood as showing concern for the negative impact decision-subjects experience when they are exposed to decisions based upon generalised assumptions, it could reach even further: could an anti-stereotyping inspired approach be used to guard people against their being judged on the basis of unreasonable or unjustifiable algorithmic generalisations and assumptions at large (including for instance the creation of new stereotypes)? [544]

---

[542] Concurring Opinion of Judge Turković in ECtHR Grand Chamber (24 January 2017) Case of Khamtokhu and Aksenchick v. Russia, Application nos. 60367/08 and 961/11, § 3.

[543] See Chapter 6, Section 6.2 "Algorithmic Stereotyping and Generalisations." This could be achieved through a set of reflective exercises, which necessitate decision-makers to lay bare the assumptions they may have relied upon: Miceli et al., 'Documenting Computer Vision Datasets'.

[544] This issue will be further explored in Chapter 6, Section Algorithmic Stereotyping and Deinvidualisation. Here, I will investigate whether algorithmic generalisations could impose socio-relational and distributional egalitarian harms (similar to stereotypes) onto decision-subjects following the use of (untruthful) algorithmic assumptions.

Vulnerability

The Court's shift towards more socially oriented conceptions of equality has also been exemplified by the introduction of the notion of (group) vulnerability.[545] A good example of the ECtHR's vulnerability rationale can be found in the case of *N. v. Romania*. In this case the ECtHR accentuated the precarious position of the mentally disabled:

> "If a restriction on fundamental rights applies to someone belonging to a particularly vulnerable group in society that has suffered considerable discrimination in the past, such as the mentally disabled, then the State's margin of appreciation is substantially narrower and it must have very weighty reasons for the restrictions in question. The reason for this approach, which questions certain classifications per se, is that such groups were historically subject to prejudice with lasting consequences, resulting in their social exclusion. Such prejudice could entail legislative stereotyping which prohibits the individualised evaluation of their capacities and needs."[546]

Relying on the notion of vulnerability, the ECtHR guards so-called vulnerable groups against the socio-relational and distributional disadvantage they have faced historically, systemically and structurally.[547] Following the ECtHR's wording, vulnerability and inequality appear as mutually

---

[545] For an exploration on the introduction of vulnerability, see also: L. Peroni and A. Timmer, 'Vulnerable Groups: The Promise of an Emerging Concept in European Human Rights Convention Law', *International Journal of Constitutional Law* 11, no. 4 (1 October 2013): 1056–85, https://doi.org/10.1093/icon/mot042. Arnardóttir, 'Vulnerability under Article 14 of the European Convention on Human Rights'. Timmer, *Strengthening the Equality Analysis of the European Court of Human Rights: The Potential of the Concepts of Stereotyping and Vulnerability.* The notion was first introduced in the Chapman v. United Kingdom Case. Here, the Court observed how States should grant special consideration to vulnerable groups, *in casu* Gypsies, during regulatory planning and decision-making processes. The notion was first introduced in the Chapman v. United Kingdom Case. ECtHR Grand Chamber (18 January 2001) Case of Chapman v. The United Kingdom (Application no. 27238/95), § 96.

[546] ECtHR Fourth Section (16 November 2021, final 16 February 2022), Case of N. v. Romania (No. 2) Application no. 38048/18, § 55. Having found a violation of Article 8, the ECtHR did not believe it necessary to examine admissibility and merits of the complaints under Article 14. In their dissenting opinion, Judge Motoc did analyse the question from this perspective concluding that there would have been a violation of Article 14. The latter case repeats the wording of the ECtHR in Kiyutin v. Russia, which concerned those who suffer from HIV and AIDS. ECtHR First Section (10 March 2011, Final 15 September 2011) Case of Kiyutin v. Russia, Application no. 2700/10, § 63. See also: ECtHR Second Section (20 May 2010, Final 20 August 2010), Case of Alajos Kiss v. Hungary (Application no. 38832/06), § 42.

[547] The idea that special consideration is to be given to this position during legislative processes indicates the ascription of a transformative component, whereby equality is seen as a driver towards change. See also: Fredman, 'Emerging from the Shadows: Substantive Equality and Article 14 of the European Convention on Human Rights'.

reinforcing phenomena: whereas a given group's vulnerability arises out of the inequalities and discrimination they have experienced in the past, due to their vulnerable position these groups are also more likely to become the subject of even greater inequality. Vulnerability acts as a type of socio-relational status and those who carry the status (risk to) face egalitarian, such as exclusion, as well dignitarian harms.

The ECtHR has also recognised how stereotypes and vulnerability are interconnected phenomena and mutually reinforcing social (status) harms.[548] Timmer believes both notions have a signalling function. On the one hand, the identification of pervasive stereotypes could lead to the identification of vulnerable groups. On the other hand, the recognition of a group's vulnerable status can signal those instances where the application of stereotypes (or generalised assumptions) is likely to be particularly harmful.

VULNERABLE GROUPS – According to Peroni and Timmer, the ECtHR developed an understanding of vulnerability that is relational, particular and harm-based.[549] It is relational as the Court situates vulnerability within a wider social context, co-shaped by social, historical and institutional forces, which act to maintain the vulnerable status of a group of which individuals are a part.[550] It is particular as the Court does not view people as inherently vulnerable. Instead, "vulnerability is shaped by specific group-based experiences" and tied to group membership.[551] Concerning the types of disadvantage vulnerable groups experience, they believe two broad categories of harm can be discerned: misrecognition and maldistribution. In "misrecognition" cases, vulnerability originates from the negative social attitudes that exist towards a certain group, such as prejudice and stigmatisation. Groups and group members are not recognised as equals, and therefore not

[548] Timmer, *Strengthening the Equality Analysis of the European Court of Human Rights: The Potential of the Concepts of Stereotyping and Vulnerability.*, 246–48. In their joint dissenting opinion to the case of *Carvalho Pinto de Sousa Morais v. Portuga,* Judges Ravarani and Bošnjak note the relationship between stereotypes and vulnerability: "it is important to note that in all cases where the Court has found discrimination based on stereotypes, there has always been a *direct* allusion to the person's belonging to a certain particularly vulnerable group in society." ECtHR (25 July 2017, Final 25 October 2017) Case of Carvalho Pinto de Sousa Morais v. Portugal, Application no. 17484/15.

[549] Peroni and Timmer, 'Vulnerable Groups', 1063–65.

[550] Peroni and Timmer, 1064.

[551] Peroni and Timmer, 1064. Put differently, Arnardóttir observes how vulnerability is not as much created by the innate characteristics of the individual, but rather through the social structures an individual finds themselves in. Arnardóttir, 'Vulnerability under Article 14 of the European Convention on Human Rights', 164.

treated as equals.[552] In "maldistribution" cases, the Court links vulnerability to the material deprivation of resources necessary for people to be able to interact with others as peers, and the social disadvantage that follows thereout.[553]

While vulnerability has often been invoked in relation to specific group-based characteristics, such as gender, sex, sexual orientation, ethnicity and disability, and can therefore be seen as a further conceptualisation and continuation of the suspect grounds paradigm, it also issues a clear departure from a formal understanding of differentiation.[554] Vulnerability signals the need for asymmetrical, rather than formal and symmetrical protection: it is not the nature of the trait that signals vulnerability; instead vulnerability signals the disadvantage, experiences and consequence groups have faced due to their possession of a particular trait.[555] The ECtHR's development of the notion vulnerability has therefore been defined by Arnardottir as a socio-contextual "identity plus approach," which remains "anchored in one or more group-defining identity markers (i.e., suspect discrimination grounds), but allows the Court to target the more complex and deeply embedded situations of disadvantage by looking towards social context."[556]

People's vulnerability-status is shaped through social processes, and not every social group is a vulnerable group.[557] Timmer believes several reasons can be put forward when labelling a group

---

[552] Of course, misrecognition can result in maldistribution: ECtHR Grand Chamber (13 November 2007) Case of D.H. and Others v. The Czech Republic, Application no. 57325/00; ECtHR Second Section (29 January 2013, Final 29 April 2013), Case of Horváth and Kiss v. Hungary, App. No. 11146/11.

[553] Peroni and Timmer, 'Vulnerable Groups', 1065. In reference to the work of Fraser. See for instance, the case of D.H. and Others v. The Czech Republic, which concerned the distribution of educational opportunities. ECtHR Grand Chamber (13 November 2007) Case of D.H. and Others v. The Czech Republic, Application no. 57325/00, § 193.

[554] See also Arnardóttir, 'Vulnerability under Article 14 of the European Convention on Human Rights', 165. However, not all vulnerable groups are defined by a "suspect" characteristic. For example, "asylum-seekers are also considered a member of a particularly underprivileged and vulnerable population group. See: ECtHR Grand Chamber (21 January 2011) Case of M.S.S. v. Belgium and Greece, Application no. 30696/09 and ECtHR Grand Chamber (16 March 2010), Case of Oršuš and Others v. Croatia, Application no. 15766/03, § 147.

[555] Arnardóttir, 166.

[556] Arnardóttir, 170.

[557] Peroni and Timmer, 'Vulnerable Groups', 1064–65. However, Kim has observed how the vulnerability concept has not always been consistently applied, and hence, identification of groups as vulnerable might at times appear selective For instance, in the case of *Carvalho Pinto de Sousa Morais v. Portugal*, the ECtHR did not (explicitly) apply the vulnerability criterion, even though the subject matter, gender violence, would historically warrant the use thereof. Only in the joint dissenting opinion of Judges Ravarani and Bošnjak a reference to the notion of vulnerability could be found. According to Kim, the ECtHR's inconsistent use of the notion may be due to their reluctance to define vulnerable groups, and instead rely upon vulnerability markers, such as social disadvantage, state dependence, prejudice and stigma. While this does allow for an expansive interpretation of the concept, the contours of the notion have remained

as vulnerable. For example, some groups, such as children and the mentally disabled, are inherently vulnerable, whereas others may be considered vulnerable because they hold a precarious relationship vis-a-vis the state (or those in power), such as prisoners in detention.[558] Vulnerability can also be linked to the absence of power to change the disadvantage one faces. Powerful persons and groups have a greater capacity to disadvantage those who are vulnerable.[559] Those identified as vulnerable therefore share an increased risk to suffer from the aforementioned relational (recognition) and distributive (misdistribution) harms. Vulnerability can thus have a signalling function to designate those sub-groups that are most disadvantaged. This in turn could let policy makers know whom they should prioritise in the distribution of important social and socio-economic resources.[560]

At the same time, the recognition of a person or group as vulnerable must be met with a certain degree of caution as the vulnerability-label itself could be experienced as stigmatising or give rise to paternalist reasoning.[561] Likewise, one should not overemphasise group-based experiences at the cost of diminishing or misrecognising the particular or different vulnerabilities individual group members may experience.[562]

---

vague as a consequence. Kim, 'Les Vulnérables', 626. ECtHR (25 July 2017, Final 25 October 2017) Case of Carvalho Pinto de Sousa Morais v. Portugal, Application no. 17484/15.

[558] Timmer, *Strengthening the Equality Analysis of the European Court of Human Rights: The Potential of the Concepts of Stereotyping and Vulnerability.*, 236. See also: Alexandra Timmer, 'A Quiet Revolution: Vulnerability in the European Court of Human Rights', in *Vulnerability: Reflections on a New Ethical Foundation for Law and Politics*, ed. Martha Fineman and Anna Grear (Farnham: Ashgate, 2013), 147–70, https://lib.ugent.be/en/catalog/pug01:4235269.

[559] However, this presence of dominance or power in relation to vulnerability is not absolute, certain groups may be vulnerable in one sense, but powerful in another. For example, the elderly may have certain needs physically or health-wise, but they may still comprise a politically powerful group. Francesca Ippolito and Sara Iglesias Sánchez, 'Introduction', in *Protecting Vulnerable Groups : The European Human Rights Framework*, ed. Francesca Ippolito and Sara Iglesias Sánchez, 1st ed., Modern Studies in European Law (London: Hart Publishing, 2015), 7, http://www.bloomsburycollections.com/book///.

[560] Arnardóttir, 'Vulnerability under Article 14 of the European Convention on Human Rights', 169. Vulnerability may thus be used to recognize the particular nature of intersectional forms of discrimination, without having to develop an intersectional approach towards discrimination. For instance, in B.S. v. Spain, the ECtHR Court observed how the Spanish courts failed to take into account the particular vulnerable status inherent to the applicant's position as an African woman working as a prostitute. ECtHR (24 July 2012, Final 24 October 2012) Case of B.S. v. Spain, Application no. 47159/08, § 62. Arnardóttir, 163–64. On the guiding function of vulnerability, see in particular: Peroni and Timmer, 'Vulnerable Groups', 1083.

[561] Peroni and Timmer, 'Vulnerable Groups', 1072–73 and 1085; Kim, 'Les Vulnérables'.

[562] Peroni and Timmer, 'Vulnerable Groups', 1072–73 and 1085; Kim, 'Les Vulnérables'.

VULNERABILITY AND ALGORITHMS – The vulnerability approach accentuates the precarious social position certain groups and people find themselves in due to their misrecognition (including misrecognition as a result of stereotyped reasoning) or maldistribution. Transposed onto the algorithmic environment, once a social group has been identified as vulnerable, algorithmic decision-makers should pay particular attention to, or prioritise the needs of, those most disadvantaged and marginalised. The former should guarantee the inclusion of the latter's interests from cradle to grave: are the interests of those in a vulnerable position sufficiently considered both during the design and deployment of algorithmic decision-making systems?

The policy objective to equalise the unequal socio-relational and material status of those who find themselves in a vulnerable position imposes upon decision-makers a duty of avoidance and transformation. On the one hand, algorithmic decisions should not render certain individuals and groups more vulnerable. They should not amplify or reinforce the prejudice or disadvantage they experience. On the other hand, decision-makers should strive towards the active inclusion of vulnerable groups. Only when these groups' interests and needs are heard and recognised, can their current situation of disadvantage be improved upon. The vulnerability rationale thus places additional responsibility on decision-makers in power, as they have a greater capacity to affect the lives of those less fortunate. [563]

Due to the increased presence of various power asymmetries within the algorithmic environment, the idea that algorithmic decision-making may render decision-subjects more "vulnerable" is plausible. Furthermore, status-harms associated with vulnerability could be experienced by individuals social groups and non-social groups. Taking into consideration however that one should be careful to expand the notion of "vulnerability" to any person or group subject to algorithmic decision-making systems, a theory of vulnerability, and more specifically, a concrete investigation into the potential socio-relational and distributional harms imposed and perpetuated by algorithms – and their impact on person's vulnerability status - would be warranted. This seems particularly true if vulnerability as an evaluative tool would gain independent meaning, i.e., not as a notion strictly tied to suspect grounds or historically disadvantaged groups, but as a socio-relational status that a society would rather not see imposed onto others, regardless of their current social make-up.[564]

---

[563] Danisi argues that the recognition of power differentials, namely the protection of minority rights against a dominant majority, may constitute the 'true function' of the ECHR. C. Danisi, 'How Far Can the European Court of Human Rights Go in the Fight against Discrimination? Defining New Standards in Its Nondiscrimination Jurisprudence', *International Journal of Constitutional Law* 9, no. 3–4 (1 October 2011): 793–807, https://doi.org/10.1093/icon/mor044.

[564] This possibility will be further explored in Chapter 6 "Socio-relational Conceptualisations of Equality"

Proportionality as a Fail-Safe?

The contextual factors identified under the margin of appreciation and very weighty reasons doctrines can function as a guide towards the identification of the Court's conceptualisation of equality. Still, a standard, open-ended proportionality review remains a valuable tool in the evaluation of (algorithmic) differentiation acts.[565] Proportionality could be seen as a broad judicial mechanism that can resolve conflicts between the interests pursued by (algorithmic) decision-making processes on the one hand, and the egalitarian harms algorithms might impose on the other hand.[566]

Two points of caution are nonetheless in order. First, in order to maintain legal certainty, avoid the oft-heard critique of judicial activism, and provide actual guidance, the various (interrelated) steps that comprise the proportionality test would need to be applied in a clearly structured and consistent manner.[567] Second, if a differentiation ground or decision is data-driven, this does not mean that their use is rational, reasonable or objectively justifiable. Whether it was, can only be (reliably) assessed given sufficient insight into the algorithmic eco-system. Put differently, the application of a proportionality test necessitates algorithmic transparency, explicability, and accountability on behalf of the decision-maker. Only then can a proper assessment be made regarding the necessity, suitability, and well-balanced nature of a differentiating measure.[568]

4.4 Conclusion

In this Chapter, I uncovered and evaluated the ECtHR's legal conceptualisation of equality in light of the egalitarian harms algorithms may produce. The ECtHR has played an authoritative role in the development and application of key concepts and doctrines related to the societal function of equality and non-discrimination. Whereas the ECtHR's conceptualisation of equality could be seen

---

[565] As noted by Baker, even "non-suspect classifications must satisfy proportionality, even if there exist rational, legitimate grounds for distinct treatment in the abstract." Baker, 'Proportional, Not Strict, Scrutiny', 886.

[566] Moller, 'Proportionality', 710.

[567] Of course, within this structure, the contextual factors that limit the margin of appreciation, or require very weighty reasons could still have a place. In this particular context, the idea that a test must be consistently applied does not mean that one should always have the same outcome. For instance, a given balancing act may have led to a different result in the past. However, the idea that a balancing of interests is executed, remains consistent. Different outcomes are therefore possible as long as the reasoning thereto has been made sufficiently explicit.

[568] See also Chapter 5, Section 5.3 "Rationality, Arbitrariness and Reliability", as well as Chapter 8, Section 8.2 "Main Research Findings."

as a mirror for equality's current status among (a broader set of) European States as a result of the consensus approach, the Court has also tried to uncover the objective or moral truth underlying the Convention in ways that do not directly rely upon the identification of a commonly accepted standard.[569] Indeed, the Convention remains a living document, whereby equality can be conceptualised in a dynamic and evolutive way. Still, would the Court's current approach towards equality – if it were to serve as an inspiration - be sufficiently robust when confronted with the challenges posed by algorithmic decisions?

THE DEFINITION OF EQUALITY AND NON-DISCRIMINATION – Over the years, the ECtHR has maintained divergent and at times conflicting definitions as to what forms of differential treatment could constitute discrimination. In recent rulings, the ECtHR has settled on an in-between definition: not all differences in treatment can amount to discrimination, but only those that are based on an identifiable, objective, or personal characteristic or "status" by which persons or groups of persons are distinguishable from one another.[570] The notion "status" should be given a wide meaning. Hence protection extends to personal characteristics that are not innate or inherent. Though this leaves the notion of discrimination quite broad and flexible, it might preclude certain algorithmically generated discrimination from finding protection. For instance, the condition for a differentiation ground to be identifiable or personal, might not correspond to algorithmically generated profiles that are opaque, dynamic, complex, (seemingly) random or non-personal in nature. Although the ECtHR has (implicitly) recognised how discrimination can take on more complex or intersectional forms, conceptual uncertainty therefore remains as to whether the ECtHR's definition of discrimination befits the realities of the algorithmic environment. In light of technology's transformative impact, and in order to maintain the practical and effective, rather than theoretical and illusory, application of the Convention, such clarification would be desirable. Full protection could be offered if algorithmic differentiation grounds are viewed as personal by virtue of the fact that they are applied onto people or have been derived through people's information.

THE JUSTIFICATION OF DISCRIMINATION – In principle, the ECtHR's standard justification approach follows a traditional three-step procedure. First, there must be a difference in treatment of persons in analogous or relevantly similar situations (or vice versa). Second, differential

[569] See : Letsas, 'Strasbourg's Interpretive Ethic', 531. Letsas refers among others to the Case of Dudgeon v United Kingdom. ECtHR (22 October 1981) Case of Dudgeon v. The United Kingdom, Application no. 7525/76, § 60.

[570] As aforementioned, sometimes the notion "personal" is omitted: "only differences in treatment based on an identifiable characteristic, or "status", are capable of amounting to discrimination within the meaning of Article 14." See ECtHR (4 May 2021) Case of Jurčić v. Croatia, Application no. 54711/15, § 63.

treatment must be objectively justified, that is, pursue a legitimate aim and have a reasonable relationship of proportionality between the means employed and the aims pursued by a measure. While the Court has not always been consistent in the application of this structure, or in the interpretation of its individual components, the ECtHR's interpretation of equality is defined by a variety of contextual factors and rationales that influence how the interests and discretionary power of decision-makers will be weighed against those of right-holders.

The ECtHR has established a pluralist, rather than a unified approach vis-à-vis equality and non-discrimination. An initially formal and procedural approach towards equality has been complemented with a vision that is more substantive in nature.[571] From a substantive perspective, due to an increased focus on core notions related equality (e.g. dignity and autonomy), as well as contextual and social indicators of inequality (e.g. existence of prejudice, stereotypes and stigma and the vulnerable position certain social groups find themselves in), the ECtHR has been able to enhance the protection the ECHR can offer against a variety of socio-relational and even distributional egalitarian harms. These harms, which can take place at a group-level, have been more difficult to address through a purely formal understanding of equality. The latter tend to focus on the nature of the traits involved (e.g., whether they are arbitrary, irrelevant, or immutable), rather than the disadvantage a given differentiating measure might impose. The formal requirement of comparability faces similar difficulties in light of more complex and socio-contextual forms of discrimination. These formal constructs not only fail to articulate the inequalities experienced by historically disadvantaged groups, or people who find themselves multiply burdened, they are also not easily mapped onto algorithmic differentiation grounds, and

[571] However, problems concerning consistency seem inherently tied to the interpretative approach of the ECHR. First, and as aforementioned, the ECHR is seen as a living instrument, and as such, it will be subject to change. Second, the ECtHR's is developed through case-law, and as such subject to differing judicial opinions. Hence, to some degree, a certain lack of coherence and consistency is to be expected, and this can also be for the good. Still, the ECHR has been criticized for this lack of guidance and explanations in their approach, because it makes guidance in terms of what is means to be equals in light of Article 14 rather difficult. See also: Letsas, 'Strasbourg's Interpretive Ethic'. Nilsson, 'Same, Same but Different', 144. Arnardottir has pointed out that the assumption that there can be some sort of unifying theory underlying non-discrimination law might be misguided. Instead, a pluralist understanding of the ECtHR's case-law might need to be embraced: "Different elements of the principle of equality are simply appropriate to different situations. Article 14 can, and does, encompass elements of the notions of 'equality before the law', 'equal protection of the law', 'equal opportunity', 'equal treatment' and 'equal results' alike." Arnardottir, 'The Differences That Make a Difference', 669–70. Similarly, Letsas too argues that with regard to the ECHR in general "there is 'no one-size-fits-all' theory of human rights because human rights play different normative roles depending on the institutional setting under which they operate." George Letsas, 'Human Rights, Legality, and the ECHR', in *A Theory of Interpretation of the European Convention on Human Rights* (Oxford: Oxford University Press, 2007), 17–36, https://doi.org/10.1093/acprof:oso/9780199203437.003.0002.

the harms the latter produce. At the same time, it is uncertain whether non-representational algorithmic aggregates would have their disadvantage recognised under the identified substantive rationales. Could non-representational algorithmic groups experience harms similar to those imposed by stereotypes, or could they ever be considered vulnerable? A clear articulation of the specific harms these decision-subjects can experience in practice would further facilitate the evaluation of algorithmic decision-making systems.

Despite inconsistencies and conceptual uncertainty, the broad definition given to discrimination, as well as a willingness to engage into a more harms and disadvantage-based reasoning, make the ECtHR's interpretation of equality a potential basis for the evaluation of algorithmically generated injustice. In this context, a formal and a substantive approach towards equality can co-exist. An open-ended and procedural model of equality can offer a gateway for the evaluation of more complex harms that arise in cases where differentiation is imposed within a dynamic, socio(-technical) environment. This also maintains the possibility for the ECtHR to evaluate classic instances of formal equality before the law and to perform an institutional check against arbitrary or random decision-making.[572] Under this hybrid view, equality would remain a site for balancing of interests, where next to the differentiation ground in question (and how this ground was produced), the socio-relational and distributional harms decision-subjects experience can be weighed against the interests of decision-makers within a proportionality review. When the equality review is applied in a transparent and consistent manner, the risk of judicial activism could be further reduced. Of course, the idea that equality should be conceptualised by the Court to cover (algorithmic) socio-relational and distributional harms is not shared by everyone. Letsas for example has observed how one should not "inflate the concept of human rights so much that it covers the whole realm of justice."[573] Within the algorithmic environment however, judicial inaction might not be neutral either.

---

[572] See on this point Arnardottir, 'The Differences That Make a Difference', 669–70. See also Chapter 5, Section 5.3.3 "Rationality, Arbitrariness and Reliability."

[573] Letsas, A Theory of Interpretation of the European Convention on Human Rights, 129. For a more in-depth investigation on how the European Convention may be interpreted see also: Letsas, A Theory of Interpretation of the European Convention on Human Rights. "One often gets the impression that the European Court of Human Rights is willing to promote interests that are normally served by social and economic human rights (as political goals) but that can textually be read into the civil rights of the ECHR. This approach would be unjustified as it cannot be based on a moral principle whose benefit the Court extends to all. The Court's institutional responsibility is to interpret civil and political rights, not social and economic rights; the latter are governed by separate normative principles and form the subject matter of separate international treaties and monitoring bodies. Piecemeal and random promotion of welfare interests is not a mark of judicial innovation. It is a cause for concern over how seriously the Court takes the values of equality and legality." Letsas, 130. At the same time, Letsas did consider elsewhere that the Court is justified in

5 Equality and Non-Discrimination in the European Union

In this Chapter, I examine the role and function of equality and non-discrimination within the European Union through an analysis of the European Charter of Fundamental Rights of the European Union (CFR), the EU Equality Directives and the case-law of the European Court of Justice of the European Union in their interpretation of the former (CJEU). Following the same core structure as Chapter 4, the Chapter will start off with a brief overview of the equality and non-discrimination provisions that will be the subject of my analysis (Section 5.1). In a second phase, and in order to uncover the EU conceptualisation of equality, I investigate those provisions through the lens of the two areas of inquiry identified within Chapter 3. First, this Chapter identifies to what extent the EU legal system maintains a closed or open-ended approach towards the notion discrimination (Section 5.2). Following this analysis, an assessment can be made whether algorithmic groups (and their group members) would face *a priori* limitations in having their disadvantaged redressed under an EU(-inspired) notion of discrimination.[574] Second, I look into the conditions under which differential treatment can be justified (Section 5.3). Through an investigation of the review processes and justification procedure, I clarify the egalitarian objectives pursued within the EU legal order, the egalitarian harms people are protected against, and the conditions under which these egalitarian interests can be overridden by decision-makers.[575] These findings are then positioned against the algorithmic environment in order to assess to whether the EU conceptualisation of equality is conceptually robust or sound to tackle the egalitarian harms algorithms risk imposing onto decision-subjects.

5.1 The Legal Framework

The EU legal order presents a multi-layered and complex legal patchwork in which the legal principle of equality has long held a foundational status.[576] Equality (and non-discrimination)

identifying "the moral values which normatively constrain the projects that states pursue on the international plane." Letsas, 'Strasbourg's Interpretive Ethic'.

[574] It is therefore related to the "Equality of Whom?" debate.

[575] Or what safeguards a decision-maker should implement to

[576] For an overview of the various functions equality has held, and their history, see: Muir, 'The Essence of the Fundamental Right to Equal Treatment'. Raphaele Xenidis, 'Shaking the Normative Foundations of EU Equality Law: Evolution and Hierarchy between Market Integration and Human Rights Rationales', 2017,

appear both as a general principle of EU law, as well as a fundamental right. As an integral component of EU primary law, the principle of equality has been enshrined within the Treaty on the EU (TEU), the Treaty on the Functioning of the EU (TFEU), and the EU Charter of Fundamental Rights (CFR). Yet, as an omnipresent principle, equality performs a variety of distinct functions, and may represent different legal traditions.[577] As a general principle of EU law, equality at times functions as a traditional, procedural, and constitutional benchmark. At others, it pursues more substantive, human rights-oriented objectives.[578] Furthermore, equality has found expression through a series of specific equality and non-discrimination laws, and most notably, the EU Equality Directives.[579]

FROM AN ECONOMIC PRINCIPLE TO A FUNDAMENTAL RIGHT – Adding to the legal and political complexity, is equality's evolution from an economically oriented notion to that of a human right.[580] The original purpose of non-discrimination law within the EU was focused on market integration.[581] Equality therefore initially appeared as tool towards this integration, rather than a principle aimed at the protection of moral or normative values.[582] Over the years though, equality

---

https://www.research.ed.ac.uk/en/publications/shaking-the-normative-foundations-of-eu-equality-law-evolution-an.

[577] Muir, 'The Essence of the Fundamental Right to Equal Treatment'. See also Chapter 3, Section 3.1.1 "Models of Equality."

[578] Muir.

[579] References to equality and non-discrimination are however present in a variety of EU acts, and not only the four equality Directives. See for example Bell and Waddington's analysis of the Work-life Balance Directive. Bell and Waddington, 'Similar Yet Different: The Work-Life Balance Directive and the Expanding Frontiers of EU Non-Discrimination Law.'

[580] Whether or not a principle can also act as a fundamental right is something I shall not discuss here.

[581] For a history of this development, see also: Sacha Prechal, 'Non-Discrimination Does Not Fall Down From Heaven: The Context and Evolution of Non-Discrimination in EU Law', *Eric Stein Working Paper*, Eric Stein Working Paper, 2009, no. 4 (2009): 2. Mark Bell, 'The Transformation of EU Anti-Discrimination Law', in *Anti-Discrimination Law and the European Union* (Oxford: Oxford University Press, 2002), https://doi.org/10.1093/acprof:oso/9780199244508.003.0008; Muir, *EU Equality Law: The First Fundamental Rights Policy of the EU*, 2018; Emma Lantschner, 'The Development of the EU Non-Discrimination Regime', in *Reflexive Governance in EU Equality Law* (Oxford: Oxford University Press, 2021), https://doi.org/10.1093/oso/9780192843371.003.0002. Bruno de Witte, 'From a "Common Principle of Equality" to "European Antidiscrimination Law"', *American Behavioral Scientist* 53, no. 12 (August 2010): 1715–30, https://doi.org/10.1177/0002764210368093.

[582] See also: Uladzislau Belavusau and Kristin Henrard, 'The Impact of the 2000 Equality Directives on EU Anti-Discrimination Law : Achievements and Pitfalls', in *EU Anti-Discrimination Law Beyond Gender*, ed. Uladzislau Belavusau and Kristin Henrard, 1st ed. (Oxford,: Hart Publishing, 2019), 6–7, http://www.bloomsburycollections.com/book/eu-

and non-discrimination evolved into a fundamental right, whereby social and moral concerns, including the protection of human dignity, autonomy and personality, gained more prominence.[583] The main exponent of this evolution was the adoption of the EU Equality Directives. The latter expanded the protection against discrimination significantly."[584] Still, the actual legal and normative effects of this on-going transformation are not always as clear, and EU Law may reflect both narratives (or traditions) concurrently.[585] Xenidis cautions against a scenario where the right to equality and non-discrimination becomes "fragmented and torn between two opposed, albeit not contradictory, rationales […] on the one hand, setting societal safeguards in form of non-discrimination rights is part of guaranteeing a fundamental right to equality, and on the other hand, private autonomy of market actors is a funding element of liberal capitalist societies."[586]

At the same time, the EU institutions are bound by their competences, which limits equality and non-discrimination in becoming a self-standing, fully-fledged human right in the EU legal order.[587] The EU certainly has broad powers, but these powers are still conferred, and therefore constrained. As a consequence, equality cannot simply be invoked and enforced on every occasion where it could be of use.[588]

---

anti-discrimination-law-beyond-gender/ch1-the-impact-of-the-2000-equality-directives-on-eu-anti-discrimination-law/. de Witte, 'From a "Common Principle of Equality" to "European Antidiscrimination Law"'.

583 Prechal, 'Non-Discrimination Does Not Fall Down From Heaven: The Context and Evolution of Non-Discrimination in EU Law', 13. See also: Raphaële Xenidis, 'Transforming EU Equality Law? On Disruptive Narratives and False Dichotomies', *Yearbook of European Law* 38 (1 January 2019): e2–57, https://doi.org/10.1093/yel/yey005.

584 See also: Prechal, 'Non-Discrimination Does Not Fall Down From Heaven: The Context and Evolution of Non-Discrimination in EU Law', 13.

585 Prechal, 12. On this transformation, see also: Bell, 'The Transformation of EU Anti-Discrimination Law'; Uladzislau Belavusau and Kristin Henrard, 'A Bird's Eye View on EU Anti-Discrimination Law: The Impact of the 2000 Equality Directives', *German Law Journal* 20, no. 5 (July 2019): 614–36, https://doi.org/10.1017/glj.2019.53. Mark Dawson, *The Governance of EU Fundamental Rights* (Cambridge: Cambridge University Press, 2017), https://doi.org/10.1017/9781107707436; Muir, *EU Equality Law: The First Fundamental Rights Policy of the EU*, 2018.

586 Xenidis, 'Shaking the Normative Foundations of EU Equality Law', 17. Xenidis, 1. Xenidis, 'Transforming EU Equality Law?'

587 Sacha Prechal, 'Equality of Treatment, Non-Discrimination and Social Policy: Achievements in Three Times', *Common Market Law Review*, no. 41 (2004): 551.

588 Prechal, 551.

5.1.1 The EU Equality Directives

The key legislative (and substantial normative) core of EU non-discrimination law is comprised by a specific number of Equality Directives. The EU Equality and Non-discrimination directives are seen as giving specific expression to the general principle of equal treatment as enshrined within the Charter.[589] Their legal basis can be found in Article 19 TFEU and Article 157 (3) TFEU, which enables the EU to take appropriate action to combat discrimination based on sex, racial or ethnic origin, religion or belief, disability, age or sexual orientation.[590] The Equality Directives explicitly define what acts of differentiation constitute prohibited forms of discrimination, and hence, unjustified violations of equality. Unlike the Charter provisions, which are phrased in an open-ended manner, the Directives have a clearly delineated scope. They govern the use of specific differentiation grounds in relation to a specific set of prized (or justice-relevant) goods. They are:

- Directive 2000/43/EC: combatting discrimination on the grounds of race and ethnic origin.[591]

---

[589] In the case of Chez, the CJEU recognised for example how the Racial Equality Directive is an expression of the principle of equality recognised within Article 21 CFR. CJEU (16 July 2015) Case C-83/14 Chez, § 42. Hanna Eklund and Claire Kilpatrick, 'Article 21 EU Charter of Fundamental Rights', in *The EU Charter of Fundamental Rights: A Commentary*, ed. S Peers et al. (Hart Publishing, 2021), 33. Although a link with Article 20 may also be discerned (especially if Article 21 is seen as giving specific expression to Article 20). See also: Mark Bell, 'Article 20: Equality Before the Law', in *The EU Charter of Fundamental Rights: A Commentary*, ed. Steve Peers et al. (Hart Publishing, 2014), 564.

[590] Article 19 TFEU allows the European Parliament and Council to adopt appropriate actions against discrimination based on sex, racial or ethnic origin, religion or belief, disability, age, or sexual orientation. Another important legal basis for combatting discrimination within the EU is Article 18 TFEU, which prohibits discrimination on grounds of nationality, and to adopt rules designed to prohibit such discrimination. Article 19 TFEU. Article 157 (3) allows the adoption of measures to ensure the equal treatment of men and women in matters of employment and occupation, including the principle for equal work or work of equal value. On the Directives and their basis and relationship with EU primary law, see also: Eklund and Kilpatrick, 'Article 21 EU Charter of Fundamental Rights', 3; Muir, 'The Essence of the Fundamental Right to Equal Treatment', 825–28; Sophia Koukoulis-Spiliotopoulos, 'The Amended Equal Treatment Directive (2002/73): An Expression of Constitutional Principles / Fundamental Rights', *Maastricht Journal of European and Comparative Law* 12, no. 4 (December 2005): 327–68, https://doi.org/10.1177/1023263X0501200402.

[591] Council Directive 2000/43/EC of 29 June 2000 implementing the principle of equal treatment between persons irrespective of racial or ethnic origin (Racial Equality Directive); Available via: https://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=CELEX:32000L0043:en:HTML

- Directive 2000/78/EC: combatting discrimination at work on the grounds of religion, belief, disability, age, and sexual orientation.[592]
- Directive 2006/54/EC: ensuring equal treatment for men and women in matters of employment and occupation.[593]
- Directive 2004/113/EC: ensuring equal treatment for men and women in the access to and supply of goods and service.[594]
- In addition to the aforementioned Directives, a final Directive was proposed but never adopted. This Directive would have combatted discrimination based upon age, disability, sexual orientation and religion or belief beyond the workplace.[595]

5.1.2 The EU Charter of Fundamental Rights

As a fundamental right enshrined within the Charter, the principle of equality provides "benchmark for judicial review and serves as a source of inspiration for the dynamic interpretation of lower instruments."[596] Within the Charter, the general principle of equality (and non-discrimination) takes on several manifestations, and alongside these manifestations, is said to relate to different legal traditions. Two provisions will be key for the purpose of this chapter: Articles 20 and 21 (1) CFR.[597]

Article 20 CFR establishes equality before the law, and simply states that

---

[592] Council Directive 2000/78/EC of 27 November 2000 establishing a general framework for equal treatment in employment and occupation; Available via: https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=celex%3A32000L0078

[593] Directive 2006/54/EC of the European Parliament and of the Council of 5 July 2006 on the implementation of the principle of equal opportunities and equal treatment of men and women in matters of employment and occupation (recast); Available via: https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A32006L0054

[594] Council Directive 2004/113/EC of 13 December 2004 implementing the principle of equal treatment between men and women in the access to and supply of goods and services; Available via: https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=celex:32004L0113

[595] Proposal for a Council Directive on implementing the principle of equal treatment between persons irrespective of religion or belief, disability, age or sexual orientation {SEC(2008) 2180} {SEC(2008) 2181}; available via: https://eur-lex.europa.eu/legal-content/en/TXT/?uri=CELEX%3A52008PC0426. See also on the reluctance against the adoption of this Directive: Xenidis, 'Transforming EU Equality Law?'

[596] Muir, 'The Essence of the Fundamental Right to Equal Treatment', 821.

[597] Beyond Article 20 and 21 (1) CFR, the Charter contains other equality provisions, aimed towards the preservation of cultural, religious and linguistic diversity equality between men and women, the rights of children, the elderly, as well as the integration of people with disabilities. Article 22-26 CFR.

Everyone is equal before the law

Article 21 (1) CFR introduces equality as a general prohibition to discrimination[598]:

> Any discrimination based on any ground such as sex, race, colour, ethnic or social origin, genetic features, language, religion, or belief, political or any other opinion, membership of a national minority, property, birth, disability, age, or sexual orientation shall be prohibited.[599]

Articles 20 and 21 (1) CFR share a close relationship, whereby Article 21 CFR is said to give particular expression to Article 20 CFR.[600] Yet in practice, their respective scope of application is not always clear.[601] Article 20 CFR is often seen as representative of equality as a general and basic principle of EU law. Here, equality appears under a standard, Aristotelian-inspired formula that

---

[598] Article 21 (2) CFR establishes the principle of non-discrimination on the basis of nationality: "within the scope of application of the Treaties and without prejudice to any of their specific provisions, any discrimination on grounds of nationality shall be prohibited." The fact that nationality discrimination is treated separately, can be seen as an indication of the distinctive contours that discrimination on grounds of nationality holds within the EU. My goal however lies in trying to uncover the general goals EU equality policies want to achieve, rather than more specific ones, and hence I will not investigate this provision separately. Eklund and Kilpatrick, 'Article 21 EU Charter of Fundamental Rights', 614.

[599] Article 21 (2) CFR states that *"Within the scope of application of the Treaty establishing the European Community and of the Treaty on European Union, and without prejudice to the special provisions of those Treaties, any discrimination on grounds of nationality shall be prohibited."*

[600] The CJEU for example has stated on several occasions that "the principle of equal treatment is a general principle of EU law, enshrined in Article 20 of the Charter, of which the principle of non-discrimination laid down in Article 21(1) CFR a particular expression." CJEU (22 May 2014), Case C-356/12 Wolfgang Glatzel v. Freistaat Bayern, ECLI:EU:C:2014:350, § 43. Angela Ward, 'The Impact of the EU Charter of Fundamental Rights on Anti-Discrimination Law: More a Whimper than a Bang?', *Cambridge Yearbook of European Legal Studies* 20 (December 2018): 33, https://doi.org/10.1017/cel.2018.11. See also CJEU (5 July 2017) Case C-190/16 Werner Fries v. Lufthansa CityLine GmbH, ECLI:EU:C:2017:513, § 29; and CJEU (29 October 2020), Case C-243/19 A v Veselības ministrija; ECLI:EU:C:2020:872, § 35.

[601] Eklund and Kilpatrick, 'Article 21 EU Charter of Fundamental Rights', 5. Moreover, sometimes the inclusion of the Charter provisions appears absent within the CJEU's case-law, even though they would be relevant to the case. Ward, 'The Impact of the EU Charter of Fundamental Rights on Anti-Discrimination Law', 43. For an example of cases where the relationship/scope between them is not clear, see: CJEU (9 March 2017) C-406/15 of Milkova; ECLI:EU:C:2017:198; CJEU (16 September 2010) Case C-149/10 Zoi Chatzi v Ypourgos Oikonomikon; ECLI:EU:C:2010:534. The latter case, which one might have expected to be dealt under Article 21 CFR, was approached via Article 20 CFR. See also: Mark Bell, 'Article 20: Equality Before the Law', in *The EU Charter of Fundamental Rights: A Commentary*, ed. Steve Peers et al. (Hart Publishing, 2021), 607.

stipulates that comparable situations should not be treated differently and different situations not be treated alike unless such treatment is objectively justified.[602] Article 21 (1) CFR is more closely related to Article 14 ECHR as both define equality as a prohibition to discriminate, whereby particular attention is paid to several, non-exhaustively listed differentiation grounds.[603]

Bell describes their relationship as one that sees Article 20 CFR as the general principle of equal treatment or *lex generalis* with a wide-ranging application, and Article 21 CFR as the *lex specialis* that is applicable only in specific situations.[604] In a similar vein, the distinction between them has been described as reflecting formal versus substantive equality.[605] Under this interpretation, which I shall maintain throughout the following sections, Article 21 CFR governs status-based inequality and differentiation imposed on the basis of "important characteristics that often give rise to inequalities across different aspects of social life."[606] Being focused on inequalities related to personal characteristics or statuses with a history of past disadvantage, Article 21 (1) CFR therefore reflects a stronger moral position against certain forms of differentiation.[607] Article 20

---

[602] See e.g. CJEU (13 April 2000) Case C-292/97, Kjell Karlsson and Others; ECLI:EU:C:2000:202; CJEU (22 May 2014), Case C-356/12 Wolfgang Glatzel v. Freistaat Bayern, ECLI:EU:C:2014:350, § 43 and 81. Xenidis observed however how the language of the CJEU evolved in relation to the general principle of equality: "From Mangold onwards, the terms used by the Court shifted from a 'principle of equality' to 'a principle of equal treatment' and then to a 'principle of non-discrimination.'" Xenidis, 'Transforming EU Equality Law?', 13.

[603] And in so far as they correspond with one another, Article 21 (1) CFR applies in compliance with Article 14 ECHR. Explanations relating to the Charter of Fundamental Rights (14 December 2007) (2007/c 303/02) OJ C303/17. Explanation on Article 20 and Article 21 CFR. In general, the Charter must be seen in light of other human rights instruments, including the ECHR. Article 53 CFR states that: "Nothing in this Charter shall be interpreted as restricting or adversely affecting human rights and fundamental freedoms as recognized, in their respective fields of application, by Union law and international law and by international agreements to which the Union, the Community or all the Member States are party, including the European Convention for the Protection of Human Rights and Fundamental Freedoms, and by the Member States' constitutions."

[604] Bell, 'Article 20: Equality Before the Law', 2014, 565.

[605] Lorena Sosa and Alexandra Timmer, 'Conceptual Underpinnings of the EU's Commitment to Human Rights', in *The European Union and Human Rights*, by Lorena Sosa and Alexandra Timmer (Oxford University Press, 2020), 62–78, https://doi.org/10.1093/oso/9780198814191.003.0004.

[606] Bell, 'Article 20: Equality Before the Law', 2014, 565. Among those aspects of social life, Bell refers to the "prized public goods" protected under the Directives, such as employment, education and access to services. Article 21 CFR may also be more concerned with differentiation upon traits that are intrinsic to one's personality. Still, this distinction is not airtight. Bell points towards the inclusion of property among Article 21's list of (explicit) protected characteristics, which is not a "stable feature of personal identity?" See also: Eklund and Kilpatrick, 'Article 21 EU Charter of Fundamental Rights'.

[607] Under a view where Article 21 (1) CFR represents a clear moral position against certain forms of status-based discrimination, cases that are covered, and dealt with under this provision, are therefore also more likely to trigger a

CFR would instead apply to "any situation where the law differentiates between categories in its application."[608] Unlike Article 21 (1) CFR, Article 20 CFR could then be considered a procedural benchmark with (less) normative content aimed to detect irrational or arbitrary forms of differentiation.[609] It would be a catch-all provision, governing those forms of (non-personal) differentiation that are not caught by Article 21 (1) CFR, but that we nonetheless want to make subject to a review.[610] In addition, Article 20 CFR mainly acts as an infrastructural or institutional

---

heightened level of scrutiny by the CJEU. See also Section 5.3: "The Justification of Differentiation." In addition, some authors have argued that 21 (1) CFR is therefore also more likely to find application in relation to the interpretation of Equality Directives, and Article 20 CFR onto provisions of equal treatment found in other branches of EU law. However, this may not always be the case, in its 2022 ruling in interpreting the Passenger Name Records Directive, the Court solely relied upon Article 21 of the Charter, rather than Article 20 CFR. Bell and Waddington, 'Similar Yet Different: The Work-Life Balance Directive and the Expanding Frontiers of EU Non-Discrimination Law.' See also: Bell, 'Article 20: Equality Before the Law', 2014, 564; Eklund and Kilpatrick, 'Article 21 EU Charter of Fundamental Rights'. See: Directive (EU) 2016/681 of the European Parliament and of the Council of 27 April 2016 on the use of passenger name record (PNR) data for the prevention, detection, investigation and prosecution of terrorist offences and serious crime, OJ L 119, 4.5.2016, p. 132–149.; CJEU (21 June 2022), Case C-817/19 Ligue des Droits Humains v. Conseils des Ministres; ECLI:EU:C:2022:491. For a discussion on the PNR ruling, see also Sections 5.2.2 "The EU Charter of Fundamental Rights" and Section 5.3.3 "Rationality, Arbitrariness and Reliability."

[608] Bell, 'Article 20: Equality Before the Law', 2014, 565. Bell, 'Article 20: Equality Before the Law', 2021, 600–601. Here, Article 20 CFR represents a tradition of case-law that typically - though not always - involves a lower-level justification test: it allows policy makers to impose and justify differential treatment on any ground as long as they can provide a good reason for doing so and that the differentiating measure is in proportion to the reason put forward. See also Section 5.3 "What Inequalities Matter? The Justification of Differentiation." Dagmar Schiek, 'Age Discrimination before the ECJ - Conceptual and Theoretical Issues', *Common Market Law Review*, no. 48 (2011): 784. Ward, 'The Impact of the EU Charter of Fundamental Rights on Anti-Discrimination Law', 41. Bell, 'Article 20: Equality Before the Law', 2021; Bell, 'Article 20: Equality Before the Law', 2014. See also Section 6.3.3. Bell, 566. See also: Muir, 'The Essence of the Fundamental Right to Equal Treatment', 829.

[609] In this regard, Article 20 CFR is similar to what has been referred to as a first-generation approach towards equality, whereby equality primary aims to secure good governmental and political governance. See also Section 4.2.2. Functions of Equality, and the models identified by Sottiaux. Sottiaux, 'Het Gelijkheidsbeginsel : Langs Oude Paden En Nieuwe Wegen (Artikel, 2008) [WorldCat.Org]'. Bell, 'Article 20: Equality Before the Law', 2014, 566. Tridimas, *The General Principles of EU Law*, 62.

[610] Such a reading may follow from the Glatzel case, where the CJEU on the one hand analysed the prohibition to discriminate on the grounds of disability under Article 21 CFR and analysed the degree at which the law respected the general principle of equal treatment under Article 20 CFR. Similarly, in the PNR case mentioned in footnote 632, the CJEU addressed the use of pre-determined criteria based upon protected characteristics through Article 21 CFR, without mention of Article 20 CFR. The CJEU does not follow this dichotomy in every case however. For instance, in the case of *Milkova*, the CJEU took a combined approach: "The principle of equal treatment is a general principle of EU law, now enshrined in Articles 20 and 21 of the Charter, which requires that comparable situations must not be treated differently and that different situations must not be treated in the same way unless such treatment is objectively justified." This merging of the general principle equality was welcomed by Muir who claimed that it could avoid relational confusion

benchmark, applicable in the vertical relationship between public institutions and citizens. Article 21 CFR is more interpersonal in nature.[611] It has been given horizontal direct effect and can find application in private relationships.[612] Yet, as far as these traditions are reflected within the Charter, these conceptions of equality co-exist and are complementary.[613]

Unfortunately, the CJEU does not always makes a clear distinction between the two provisions and there is no clear boundary which defines their scope of application.[614] Still, for the sake of

---

between the two principles and ensure coherence. CJEU (18 July 2013) Case C-356/12 of Wolfgang Glatzel v. Freistaat Bayern (Glatzel); Denis Martin, 'Article 20 CFR', in *Article 20 CFR* (Oxford University Press, 2019), 2162, https://doi.org/10.1093/oso/9780198759393.003.541. CJEU (9 March 2017) Case C-406/15 Milkova; ECLI:EU:C:2017:198, § 55. Muir, 'The Essence of the Fundamental Right to Equal Treatment', 832. . CJEU (21 June 2022), Case C-817/19 Ligue des Droits Humains v. Conseils des Ministres; ECLI:EU:C:2022:491.

[611] Of course, Article 21 CFR is also institutional in the sense that it also prohibits institutions from imposing discrimination on the basis of status-grounds. Elise Muir, 'The Transformative Function of EU Equality Law', *European Review of Private Law* 21, no. 5/6 (1 November 2013): 1233–34 and 1241., https://kluwerlawonline.com/journalarticle/European+Review+of+Private+Law/21.5/ERPL2013075.

[612] Horizontal direct effect is typically understood as the capacity for individuals to rely directly upon the Charter before the national Courts in disputes between private purposes. Thus far, the CJEU has not yet been asked to consider whether Article 20 can be directly relied upon in interpersonal disputes. Egenberger laid down the horizontal direct effect of Article 21 CFR on the basis of religion, Mangold on the basis of age. Bell, 'Article 20: Equality Before the Law', 2021, 612. CJEU (17 April 2018) Case C-414/16 Egenberger v. Evangelisches Werk Fur Diakonie und Entwicklung eV; ECLI:EU:C:2018:257, § 76 and 77: "The prohibition of all discrimination on grounds of religion or belief is mandatory as a general principle of EU law. That prohibition, which is laid down in Article 21(1) of the Charter, is sufficient in itself to confer on individuals a right which they may rely on as such in disputes between them in a field covered by EU law." See also: CJEU (22 November 2006) Case C-144/04 Werner Mangold v. Rüdiger Helm, ECLI:EU:C:2005:709, § 74-78. For more information on the horizontal direct effect of the Charter, see also: Sacha Prechal, 'Horizontal Direct Effect of the Charter of Fundamental Rights of the EU', *Revista de Derecho Comunitario Europeo*, no. 66 (n.d.): 407–26, https://doi.org/10.18042/cepc/rdce.66.04.

[613] Muir, 'The Transformative Function of EU Equality Law', 1234. See also Chapter 3, Section 3.1. "Equality: A Multi-Layered Legal Notion."

[614] Bell, 'Article 20: Equality Before the Law', 2014, 566. The distinction between Article 20 and 21 CFR should perhaps not be as sharply drawn, however. The possibility for certain forms of differentiation - even if they are not immediately "status-related" – to be captured under Article 21 CFR should therefore not be precluded. If Article 21 (1) CFR is indeed more normative (or substantive) in content, other forms of (algorithmic) differentiation could at one point in time find themselves governed through Article 21 CFR (and subject to a higher level of scrutiny). For instance, society may evolve and discover that certain forms of differentiation that were once deemed justified, suddenly are no longer justifiable. At the same time, even if Article 20 CFR would continue to apply, the CJEU may always choose to heighten its review in relation to certain forms of differentiation. See also Section 5.3. The Justification of Differentiation. In this regard, Bell refers to the observations of Advocate-General Sharpston in relation to age discrimination. Over the years, people started looking differently at the use of age as a differentiating factor. While accepted at first, age discrimination became

clarity, and in line with aforementioned analysis on the basis of Bell, I shall continue to make use of a distinction whereby Article 20 CFR is predominantly seen as representing a tradition of case-law concerning the general principle of equal treatment (as a constitutional benchmark) that applies to differences that are not closely linked to the status-based characteristics (explicitly) mentioned in Article 21.[615]

SCOPE OF APPLICATION - Both equality clauses of the Charter address Member States in their implementation of Union law.[616] Though this clearly limits their field of application, the potential reach Articles 20 and 21 CFR hold as interpretative instruments shaping the function and role of equality in the EU is substantial.[617] The EU legislator has launched several regulatory initiatives with a clear reference to fundamental rights, including the rights to equality and non-discrimination. These initiatives (will) give ground for the Charter (and Articles 20 and 21 therein) to find application as they introduce new areas that (will be) covered by EU law. Amongst them, the General Data Protection Regulation (GDPR) and the Digital Services Act and the proposed Artificial Intelligence Acts are the most notable for algorithmically guided decision-making. These instruments issue a clear mandate to decision-makers to avoid discrimination and to guarantee equality within their respective scopes.[618] Given personal data often underlie

more strictly scrutinised. Bell, 566. Opinion of AG Sharpston (22 May 2008) Case-427/06 Birgit Bartsch v Bosch und Siemens Hausgeräte (BSH) Altersfürsorge GmbH; ECLI:EU:C:2008:297. See also

[615] Bell notes however that this distinction does not give a full explanation, and hence both articles should not be mutually exclusive. Bell, 'Article 20: Equality Before the Law', 2021, 601.

[616] Article 51, § 1 CFR, as well as Article 6, §1 of the Treaty on European Union, which stipulates that "The provisions of the Charter shall not extend in any way the competences of the Union as defined in the Treaties."

[617] Muir, *EU Equality Law: The First Fundamental Rights Policy of the EU*, 2018, 201.

[618] The GDPR protects natural persons with regard to the processing of their personal data. The Recitals recognise how the processing of personal data might give rise to discrimination (see among others Recitals 71 and 75 GDPR). The proposed AI Act lays down harmonised rules on artificial intelligence. The proposal recognises how certain algorithmic process might come into conflict with the principle of equality and non-discrimination (see among others Recitals 15, 17 and 44 Proposed AIA). Finally, the Proposed Digital Services Act, which establishes a single market for digital services, also emphasises the impact certain platforms that make us of algorithms, such as recommender systems, can have on equality and non-discrimination (see among others Recital 63 Proposed DSA). In a resolution on the use of AI, the European Parliament issued a proposal that demands that "Any software, algorithm or data used or produced by high-risk artificial intelligence, robotics and related technologies developed, deployed or used in the Union shall be unbiased and [...] shall not discriminate on grounds such as race, gender, sexual orientation, pregnancy, disability, physical or genetic features, age, national minority, ethnicity or social origin, language, religion or belief, political views or civic participation, citizenship, civil or economic status, education, or criminal record." Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation) (Text with EEA relevance), available via: https://eur-lex.europa.eu/eli/reg/2016/679/oj;

algorithmically guided decision-making processes, and given the fact that algorithmically guided decision-making processes often constitute a form of artificial intelligence that has found its use in the provision of digital services, the potential applicability of the Charter's equality provisions on algorithmic differentiation and inequality within the context of these instruments should not be understated.

Special attention must be given to the CJEU, which through its case-law, and within the confines of its competences, has given conceptual shape to the principle of equality as it applies to EU law.[619] More specifically, like the ECtHR, the CJEU too has further refined the role of equality within the EU and the concept of discrimination therein. The case-law of the CJEU may thus have a considerable influence on the way equality is given shape in future EU legislation and how notions of discrimination are to be interpreted.[620]

Proposal for a Regulation of the European Parliament and of the Council Laying Down Harmonised Rules on Artificial Intelligence (Artificial Intelligence Act) and Amending Certain Union Legislative Acts (COM/2021/206 Final); available via: https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A52021PC0206. Proposal for a Regulation of the European parliament and of the Council on a Single Market for Digital Services (Digital Services Act) and amending Directive 2000/31/EC (Com/2020/825 Final); available via: https://eur-lex.europa.eu/legal-content/en/TXT/?qid=1608117147218&uri=COM%3A2020%3A825%3AFIN . Article 9 European Parliament resolution of 20 October 2020 with recommendations to the Commission on a framework of ethical aspects of artificial intelligence, robotics and related technologies (2020/2012(INL); 2021/C 404/04); available via: https://eur-lex.europa.eu/legal-content/EN/TXT/HTML/?uri=CELEX:52020IP0275

[619] Muir, 'The Transformative Function of EU Equality Law'.

[620] For instance, in a recent line of cases, the CJEU had to consider the non-discriminatory nature of differentiating criteria when processing PNR data through algorithms. Sections 5.2.2 "The EU Charter of Fundamental Rights" and Section 5.3.3 "Rationality, Arbitrariness and Reliability."

5.2 The (Initial) Range of Protection

In Chapter 3, it was established how the conceptualisation of equality within equality and non-discrimination laws could be uncovered through two areas of enquiry: an exploration into the definitional scope of the notion discrimination and an investigation into the evaluative mechanisms through, and conditions under which, differential treatment can be justified. In this Section the definitional scope of the notions of equality and non-discrimination will be examined.[621]

5.2.1 The EU Equality Directives

The EU Equality Directives grant people an (individual) right not to be discriminated against on the basis of specifically enumerated, protected characteristics in their access to specific economically, socially or culturally prized goods, such as protection against sex-based discrimination (status) in matters of employment (a prized or justice-relevant good).[622] Unlike the Charter provisions, the EU legislator clearly chose for the EU Equality Directives to follow a closed and grounds-based approach: only where a differentiating act can be (causally) linked or tied to a protected characteristic, could there be discrimination, and hence a violation of equality.[623] This

[621] In Chapter 3, the argument was made that legislators and the judiciary often make definitional choices in their interpretation of the notions of equality and discrimination. Under an open-ended approach no *a priori* limitations are placed onto the notion of discrimination: any differentiation act can amount to discrimination and be made subject to review. Because differentiation grounds are tied to algorithmic groups, any algorithmically defined group can launch a claim to have the disadvantage associated with their differential treatment evaluated. Under a closed approach, certain differentiation grounds will be excluded from the law's protective scope.

[622] The Directives prohibit direct discrimination, which occurs where an individual, or group of individuals, is subject to differential or unfavourable treatment in relation to another individual or group of individuals (that is in a comparable situation) on the basis of a specific, and "protected", characteristic; as well as indirect discrimination, which concerns neutral provisions, criteria or practices that nonetheless place people that possess or share a particular protected characteristic at a particular disadvantage compared to other groups.

[623] Stefan Sottiaux and Jogchum Vrielink, 'Verschillende Gelijkheden? Recente Europese En Nationale Rechtspraak over Discriminatie', in *Publiekrecht*, ed. Andre Alen, Gunther Maes, and Stefan Sottiaux, Themis Cahier (Brugge: Die Keure, 2012), 26. Daniel Le Métayer and Julien Le Clainche, 'From the Protection of Data to the Protection of Individuals: Extending the Application of Non-Discrimination Principles', in *European Data Protection: In Good Health?*, ed. Serge Gutwirth et al. (Dordrecht: Springer Netherlands, 2012), 325, https://doi.org/10.1007/978-94-007-2903-2_15. See also section 5.3.1 The EU Equality Directives

legal approach is the result of a political decision-making process. Seen in isolation, the EU Equality Directives advance a constrained and narrow definition of (wrongful) discrimination.[624]

	Race and Ethnicity	Sex[625]	Religion or Belief	Disability	Age	Sexual Orientation
Education	x					
Goods and Services	x	x				
Social Protection	x	x				
Employment	x	x	x	x	x	x

Table 1 The EU Equality Directives: Status plus Prized Goods Approach (Simplified Overview)[626]

The EU Equality Directives present a straightforward mechanism through which algorithmically guided decision-making practices can be evaluated. If one were to systematically map all the protected characteristics and link them to the domains in which their use is forbidden, the resulting document would provide a set of circumstances under which the deployment of algorithms should be met with more caution.[627] When developing or deploying an algorithmically guided system, decision-makers should ask themselves the following question: will my system be used in one of the areas covered by any of the Equality Directives, and if so, is it likely to disadvantage, either directly or indirectly, an individual or group of individuals on the basis of one of the characteristics listed? [628] The relative preciseness of the Equality Directives comes at a cost.

[624] In practice, the EU Equality Directives are not the only instruments where equality and non-discrimination matter however. See also frameworks in footnote 634.

[625] This also includes discrimination on the basis of sex and maternity (Article 4 (a) Gender Equality Directive) and Article 2, § 2 (c) Gender Recast Directive.

[626] See also: Xenidis, 'Shaking the Normative Foundations of EU Equality Law', 35.

[627] Indeed, as explained in Chapter 2, distributive principles, such as the ones provided by the Directives, are good at indicating domains or contexts where the deployment of algorithms must be met with a higher degree of caution because they necessarily impose differentiating treatment and thus will deviate from the default rule that everyone should be treated equally.

[628] Beyond the limited set of differentiation grounds that can constitute a violation of equality, other difficulties concerning their application might emerge. Suppose a hiring algorithm is developed by an external software developer, for which they mainly source data from companies with a staff that is predominantly male. Because women are excluded

In light of the many contexts in which algorithmic learning techniques can be deployed and the novel differentiation grounds these technologies can discover, many (potentially) unjustified algorithmically guided decision-making processes are not captured by Directives. For example, educational algorithms that predict students' scores have been subject to public outcry due to their alleged unfairness. Criticisms levelled against their use was often twofold. First, due to their reliance on historical school-level results, students were differentiated on the basis of attributes that did not reflect their individual potential and merit. Second, due to socio-economic contextual factors, their deployment often led to classist outcomes, promoting elite-school over public school students.[629] The Directives only protect against discrimination and social exclusion on the basis of race and ethnicity for the allocation of educational opportunities however. [630] They offer no protection against the use of socio-economic attributes, nor do they implement a general prohibition on the use of non-merit based factors in the field of education.

Extending the Meaning of Differentiation Grounds

Even though the Directives exhaustively enumerate what differentiation grounds will be protected, their scope of protection has been expanded. The CJEU has interpreted the Equality Directives' protective scope to capture distinct and initially non-listed differentiation grounds due to the close relationship they share with grounds that have been explicitly listed. Rather than introduce new protected grounds, the CJEU has extended the meaning of certain grounds to also encapsulate non-listed ones. For example, in the landmark *P v S and Cornwall County Council Case*, the CJEU argued that the protection of gender equality within the area of employment should not be confined to differentiation based on the fact that a person is of one or another sex.[631] Taking a

---

from the training data, women are less likely to be hired by those companies who use their application. The exclusion of a particular group from a given decision-making process could already constitute a moral and social egalitarian harm in and of itself. A few questions emerge, however. First, is the development of the hiring algorithm a practice covered by the Gender Equality in Employment Directive? Second, is the exclusion in and of itself a harm governed by the Directive? Would it constitute a less favourable treatment even if no distributional harm has been coupled to the exclusion of women? Or would the practice only appear as discriminatory if it actually resulted in the non-hiring of women?

[629] Hern, 'Ofqual's A-Level Algorithm'; Stewart, Weale, and Proctor, 'Ofqual Chief to Face MPs over Exams Fiasco and Botched Algorithm Grading'; Han, 'An Algorithm Shouldn't Decide a Student's Future'; BBC News, 'A-Levels and GCSEs'.

[630] Le Métayer and Le Clainche, 'From the Protection of Data to the Protection of Individuals', 325. For additional examples of areas not covered by the Directives, but related to the protected grounds, see also: Gerards et al., *Algorithmic Discrimination in Europe*.

[631] CJEU (30 April 1996), Case C-13/94 P v S and Cornwall County Council; ECLI:EU:C:1996:170, § 20.

teleological approach, the CJEU believed that in view of the purpose and nature of the law, the notion of discrimination should extend to cases of unfavourable treatment that arise from a person's gender reassignment, which it argued is essentially, if not exclusively, based on the sex of the person concerned.[632] Similarly, the CJEU has recognised how certain differentiating grounds and measures are "inextricably" linked to a protected characteristic and can therefore constitute discrimination.[633] By extending the meaning of a ground, the Court can guard certain groups against discrimination where they would otherwise be excluded from protection.

The CJEU can contribute to the function of equality and non-discrimination within society by purposefully interpreting the Equality Directives. When a system is constructed around a limited or closed set of protected grounds, their ability to do so is particularly useful. It allows for an evolutive approach towards equality even in situations of legislative inaction.[634] In the absence of specific non-discrimination laws, a renewed interpretation of a ground's protective scope can

---

[632] CJEU (30 April 1996), Case C-13/94 P v S and Cornwall County Council; ECLI:EU:C:1996:170, § 20-21. Moreover, the CJEU observes how such discrimination fails to respect the dignity and freedom to which a trans individual is entitled. The CJEU concluded that it was therefore prohibited to dismiss a trans individual for a reason related to a gender reassignment

[633] For example, the unfavourable treatment of women related to pregnancy or maternity constitutes direct discrimination on grounds of sex. Recital 23, Gender Recast Directive. For an overview of cases concerning the notion of "inextricably linked", reference can be made to: CJEU (26 January 2021) Case C-16/19 VL v. Szpital Kliniczny im. dra J. Babińskiego Samodzielny Publiczny Zakład Opieki Zdrowotnej w Krakowie. At the same time, the CJEU may also clarify when a neutral criterion is not inextricably linked to a protected ground. For example, the CJEU has observed that "an internal rule of an undertaking, prohibiting workers from wearing any visible sign of political, philosophical or religious beliefs in the workplace, does not constitute, with regard to workers who observe certain clothing rules based on religious precepts, direct discrimination on the grounds of religion or belief, for the purpose of that directive, provided that that rule is applied in a general and undifferentiated way." However, "A prohibition which is limited to the wearing of conspicuous, large-sized signs of political, philosophical or religious beliefs is liable to constitute direct discrimination on the grounds of religion or belief." This criterion is inextricably linked to one or more specific religions or beliefs and would place those who practice said religions subject to unjustified direct discrimination. See for instance, CJEU (15 July 2021), Cases C-804/18 and C-341/19 WABE and MH Muller Handel. Furthermore, when interpreted in a sufficiently flexible manner, the understanding that certain traits are "inextricably linked" to a protected ground could prove beneficial where it can be shown that a data-driven differentiation ground was a proxy for a protected characteristic. Instead of being considered a form of indirect discrimination, the use of the ground could be re-categorised as a direct form of discrimination. Such a re-categorisation could for instance occur where a strong correlation was found between a given data-driven profile and a protected characteristic. Those made subject to the profile would then have a better claim against the adverse treatment they experience whereas indirect discrimination can be more easily justified, direct discrimination cannot.
See: CJEU (15 July 2021) Joined Cases C-804/18 and C-341/19 IX v. WABE EV and MH Müller Handels GmbH v. MJ.

[634] See also observations in Marjolein van den Brink and Peter Dunne, 'Trans and Intersex Equality Rights in Europe: A Comparative Analysis.' (LU: European Commisson, 2018), https://data.europa.eu/doi/10.2838/75428.

broaden the number of situations that can give rise to a rightful non-discrimination claim. Unfortunately, not every ground is as flexible or malleable for it to allow new meanings to be read into them. Moreover, even if such efforts would be lauded, the Court's redefinition carries within it an added responsibility: while some groups could suddenly find protection, other groups equally deserving of protection might be left out.[635] Furthermore, the CJEU could be criticised for entering the political landscape when it decides to do so.[636] In this regard, the Court has maintained a cautious approach towards expanding the law's scope of protection.

Extending the Number of Differentiation Grounds through the Charter

The CJEU has also been reluctant to add new differentiation grounds to the EU Equality Directives. In several cases, the CJEU responded negatively to the question whether Article 20 CFR or Article 21 (1) CFR could allow for a more expansive interpretation of the EU Equality Directives.[637] In the case of *Chacón Navas,* the CJEU was asked whether the prohibition to discriminate on the grounds of religion, belief, disability, and age in the area of employment imposed by Directive 2000/78/EC could be extended to cover the unfavourable treatment of people on the ground of sickness.[638] While the CJEU acknowledged that there is a general principle of equality and non-discrimination, which is furthermore binding on Member States where they implement EU law, it does not follow that the scope of Directives can be extended by analogy to include discrimination grounds that have not been listed.[639] In the case of *Kaltoft*, the CJEU reached a similar conclusion vis-à-vis the differentiation ground "obesity," which had been used to motivate an employmee's dismissal. The CJEU noted that Article 19 TFEU, which governs the competences of the EU in relation to non-discrimination, does not refer to obesity. Likewise, EU secondary legislation does not lay down a general principle of non-discrimination on grounds of obesity in the area of employment and occupation. There was therefore no indication that a dismissal due to obesity would fall within the scope of EU law. As a consequence, the provisions of the Charter too remained inapplicable.[640]

---

[635] The redefinition of a differentiation ground should be sufficiently nuanced. For instance, van den Brink and Dunne observe how the CJEU "paints a highly medicalised picture of trans people – conceptualising their equality protections through a lens of 'gender reassignment'." As a consequence, individuals who suffer from transphobic discrimination but opted not have any physical surgery face legal uncertainty in terms of the protection they can enjoy. van den Brink and Dunne, 53–54. See also: van den Brink and Dunne, 'Trans and Intersex Equality Rights in Europe'.

[636] See also discussions concerning judicial activism in Chapters 3 and 4.

[637] See also: Xenidis, 'Transforming EU Equality Law?', 33.

[638] CJEU (11 July 2006) Case C-13/06 Sonia Chacón Navas v. Eurest Colectividades SA *(Chacón Navas)*, § 47.

[639] CJEU (11 July 2006) Case C-13/06 Sonia Chacón Navas v. Eurest Colectividades SA *(Chacón Navas)*, § 56-57.

[640] CJEU (18 December 2014) Case C-354/13 of Fag og Arbejde (FOA) v. Kommunernes Landsforening (KL) *(Kaltoft),* § 39, in reference to: CJEU Grand Chamber (26 February 2013) Case C-617/10 of Åklagaren v. Hans Åkerberg Fransson,

In both cases, the Advocate Generals' (AG) opinions provide further clarification for the CJEU's reluctance. In his opinion to the case of *Chacon Navas*, AG Geelhoed argued in favour of a restrained approach vis-à-vis the interpretation of the Directives. Above all, the CJEU should respect choices made by the legislature. First, the Directives, as well as their legal basis (Article 19 TFEU, then Article 13 EC), provide an exhaustive description of the prohibited differentiation grounds.[641] Second, the legislature has to make tragic choices in their definition of discrimination. These choices are made (among others) in an effort to balance the interests of those differentiated against the financial and economic consequences a prohibition to discriminate entails.[642] In conclusion, neither the wording, nor the history of EU law allow for an expansion of grounds.

In the case of *Kaltoft*, AG Jaaskinnen addressed the question as to whether there could be a general, open-ended prohibition on discrimination, including obesity, in the labour market. To establish a general principle of non-discrimination in EU law on all forms of differentiation would breach the established boundary of EU fundamental rights law.[643] First, EU law precludes recourse to the Charter, including Article 21(1) CFR, as a means to extend in any way the competences of the EU as defined in the treaties.[644] Second, the mere fact that discrimination takes place within the labour market is insufficient to conclude that discrimination took place during the

---

§21-22. In the Fransson case, the CJEU clarified that "since the fundamental rights guaranteed by the Charter must therefore be complied with where national legislation falls within the scope of European Union law, situations cannot exist which are covered in that way by European Union law without those fundamental rights being applicable. The applicability of European Union law entails applicability of the fundamental rights guaranteed by the Charter." However, "Where, on the other hand, a legal situation does not come within the scope of European Union law, the Court does not have jurisdiction to rule on it and any provisions of the Charter relied upon cannot, of themselves, form the basis for such jurisdiction." Obesity could in some cases be considered a disability, which is a listed ground. However, in that case, obesity must entail "a limitation resulting in particular from long-term physical, mental or psychological impairments which in interaction with various barriers may hinder the full and effective participation of the person concerned in professional life on an equal basis with other workers." CJEU (18 December 2014) Case C-354/13 of Fag og Arbejde (FOA) v. Kommunernes Landsforening (KL) *(Kaltoft),* § 59-60.

641 Opinion of AG Geelhoed (16 March 2006) in Case C-13/06 Sonia Chacón Navas v. Eurest Colectividades SA *(Chacón Navas)*, § 50-56. This exhaustive nature, Advocate General Geelhoed contrasts to the classical, open-ended approach found in human rights treates, including Article II-81 of the Constitutional Treaty (now Article 21 (1) CFR). See also: Delia Ferri, 'The Unorthodox Relationship between the EU Charter of Fundamental Rights, the UN Convention on the Rights of Persons with Disabilities and Secondary Rights in the Court of Justice Case Law on Disability Discrimination', *European Constitutional Law Review* 16, no. 2 (June 2020): 275–305, https://doi.org/10.1017/S1574019620000164.

642 Opinion of AG Geelhoed (16 March 2006) in Case C-13/06 Sonia Chacón Navas v. Eurest Colectividades SA *(Chacón Navas*, § 55.

643 Opinion of Advocate General Jääskinen (17 July 2014) Case C-354/13 Kaltoft, ECLI:EU:C:2014:2106

644 This follows from Article 6 TEU and Article 51 (2) CFR. Opinion of Advocate General Jääskinen (17 July 2014) Case C-354/13 Kaltoft, ECLI:EU:C:2014:2106, § 19.

implementation of EU law.[645] Instead, the application of EU fundamental rights law requires a certain degree of connection with EU law: "a requisite link will be established when there is a specific and identified provision of Member State law […] falling within the (substantive) scope of an equally specific and identified provision of EU law, whether it is found in an EU legislative act, or in the Treaties themselves." Third, the AG observed how all legislative acts that prohibit discrimination do so in relation to specific grounds and specific subject areas. Hence, they do not preclude in a generalised manner any discriminatory treatment.[646]

INTERSECTIONAL DISCRIMINATION - The CJEU has also been hesitant to acknowledge the particular nature of complex forms of multiple or intersectional discrimination.[647] The most notable opportunity for the CJEU to recognise the impact of intersectional discrimination concerned the case of *Parris*. Mr. Parris claimed to have been discriminated on the basis of his age and sexual orientation as he could not enjoy a survivor's pension benefit from his employer on account of the combined effects of both age and sexual orientation. [648] AG Kokott had advised the CJEU to duly take into account how "the combination of two or more different grounds for a difference of treatment is a feature which lends a new dimension to a case."[649] The CJEU decided otherwise however: "while discrimination may indeed be based on several of the grounds […], there is […] no new category of discrimination resulting from the combination of more than one of those grounds, such as sexual orientation and age, that may be found to exist where discrimination on the basis of those grounds taken in isolation has not been established."[650] Likewise, the CJEU's case law concerning the application of neutrality rules vis-à-vis the carrying of religious symbols in the work environment has been critiqued for overlooking the intersectional harms such measures

---

[645] Opinion of Advocate General Jääskinen (17 July 2014) Case C-354/13 Kaltoft, ECLI:EU:C:2014:2106, § 22.

[646] Opinion of Advocate General Jääskinen (17 July 2014) Case C-354/13 Kaltoft, ECLI:EU:C:2014:2106, § 24.

[647] See also Chapter 4, Section 4.2. "The Range of Protection"

[648] In order to enjoy the benefit, members had to either be married or entered into a civil partnership before reaching the age of 60. Unfortunately, Parris had already reached the age of 60 when this civil partnership became legally possible. CJEU (24 November 2016) Case C-443/15 of Parris v. Trinity College Dubline (and others).

[649] Advocate General Kokott thus advised the CJEU to consider the intersectional - and not only the multiple - nature of the discrimination, even referring to intersectional literature doing so. She concluded: "All things considered, therefore, indirect discrimination […] would have to be assumed to be present if it proved to be the case that the terms of the occupational pension scheme at issue do not pose a disadvantage on the basis of age alone or on the basis of sexual orientation alone, but do so on the basis of a combination of those two grounds for a difference of treatment. AG Kokott (30 June 2016) Opinion in Case C-443/15 of Parris v. Trinity College Dublin (and others), §149 – 150; 153 and 159. With regard to multiple discrimination, AG Kokott referred to the cases *Kleist, Odar and Milkova. B*oth the Racial Equality and Framework Employment Directive do recognise in their recitals that women are often the victims of multiple discrimination. Recital 14 Racial Equality Directive and Recital 3 Framework Employment Directive.

[650] CJEU (24 November 2016) Case C-443/15 of Parris v. Trinity College Dubline (and others), § 80.

impose on women wearing headscarves. In the latter cases, not only is a person's religion at play, but also their sex – and often times their ethnicity or social status as well.[651]

Even though the CJEU appears unwilling to expand the protective scope of the EU Equality Directives by adding new protective grounds, it has extended the protection of differentiation grounds *ratione personae*. That is, the CJEU recognises discrimination by association. In the landmark case of *Coleman*, the Court acknowledged that the unfavourable treatment of an employee based upon the disability of that employee's child, rather than the disability of the employee themself, could also amount to discrimination.[652] Whereas Coleman concerned direct discrimination by association, the case of *Chez* is typically seen as the CJEU's move towards the protection against indirect discrimination by association.[653] In *Chez,* the CJEU observed how measures which indirectly discriminate against an ethnic minority could also affect and damage people who do not share the ethnicity in question.[654]

---

[651] Schiek, 'On Uses, Mis-Uses and Non-Uses of Intersectionality before the Court of Justice (EU)'; Dagmar Schiek, 'Broadening the Scope and the Norms of EU Gender Equality Law: Towards a Multidimensional Conception of Equality Law', *Maastricht Journal of European and Comparative Law* 12, no. 4 (December 2005): 427–66, https://doi.org/10.1177/1023263X0501200405. Cara Donegan, 'Thinly Veiled Discrimination : Muslim Women, Intersectionality and the Hybrid Solution of Reasonable Accommodation and Proactive Measures', *European Journal of Legal Studies*, no. 2 (30 November 2020): 143–79, https://doi.org/10.2924/EJLS.2019.032. Moreover, the CJEU chose investigate these claims in relation to the ground of religion. CJEU (14 March 2017) Case C-157/15 Samira Achbita and Centrum voor Gelijkheid van Kansen en voor Racismebestrijding v. G4S Secure Solutions; CJEU (13 July 2016) Case C-188/15 Asma Bougnaiou and Association de Defense des Droits de l'Homme v. Micropole SA. CJEU (15 July 2021) and Joined Cases C-804/18 and C-341/19 IX v. WABE EV and MH Müller Handels GmbH v. MJ.

[652] CJEU (17 July 2008) Case C-303/06 Coleman, ECLI:EU:C:2008:415.

[653] CJEU (16 July 2015) Case C-83/14 Chez, ECLI:EU:C:2015:480; Benedi Lahuerta, 'Ethnic Discrimination, Discrimination by Association and the Roma Community: CHEZ'; Mariana Schafhauser Boçon, 'The CJEU's Judgment in CHEZ and Indirect Discrimination by Association: An Analysis of Progress in Combating Racial or Ethnic Discrimination in the European Union', *UNIO – EU Law Journal* 3, no. 2 (12 March 2019): 20–33, https://doi.org/10.21814/unio.3.2.3. The latter case concerned the placement of electricity meters by the Bulgarian electrical company CHEZ. In areas with a predominant Roma population electricity meters were placed on pylons at a height of six and seven meters, whereas in other districts the electricity meters were more easily accessible. According to CHEZ, the measure was taken in order to avoid tampering with the meters. Yet, if not objectively justified, it could constitute an indirectly discriminatory measure. Even though the complainant did not establish she herself was Roma, the CJEU maintained its stance in Coleman that protection is not limited to persons possessing the protected characteristic. CJEU (16 July 2015) Case C-83/14 Chez, § 56.

[654] See also Waddington's analysis of the Kaltoft case, which could have addressed the issue of discrimination against "assumed" disability. The CJEU however was not asked to rule on this legal question, nor did it bring up the issue itself. Lisa Waddington, 'Saying All the Right Things and Still Getting It Wrong: The Court of Justice's Definition of Disability and Non-Discrimination Law', *Maastricht Journal of European and Comparative Law* 22, no. 4 (August 2015): 589, https://doi.org/10.1177/1023263X1502200406. Some scholars have argued however that the Chez case can be seen

5.2.2 The EU Charter of Fundamental Rights

The CJEU has taken a cautious approach in adding new discrimination grounds to the Equality Directives. The question can nonetheless be raised whether the Charter, and Articles 20 and 21 CFR in particular, could allow for an expansive approach towards equality and non-discrimination when Member States implement EU law in areas not covered by the EU Equality Directives.

Article 20 CFR

Article 20 CFR – Article 20 CFR embodies the general principle of equality and non-discrimination under the classic formulation of equality which stipulates that comparable situation must not be treated differently, and different situations must not be treated in the same way, unless such can be objectively justified.[655] The wording of Article 20 CFR suggests that its scope of application is broad and open-ended. In his analysis of Article 20 CFR, Bell remarks that "by proclaiming 'everyone' to be equal before the law, Article 20 potentially allows anyone to mount a challenge to any difference in treatment that arises in the activities of the Union's institutions and other bodies, or within EU legislation and national implementing measures."[656] Article 20 CFR acts as a "brake on public bodies from acting in an arbitrary or inconsistent manner."[657] In other words, Article 20 CFR is an institutional and infrastructural, rather than an interpersonal benchmark. Yet, even within these confines, this Charter provision could be highly malleable. It assumes a possibility of being applied to any form of differentiation even in cases where the differentiation grounds in question have no wider social significance.

---

as moving towards this direction. Directorate-General for Justice and Consumers (European Commission) et al., *The Ongoing Evolution of the Case-Law of the Court of Justice of the European Union on Directives 2000/43/EC and 2000/78/EC: A Legal Analysis of the Situation in EU Member States* (LU: Publications Office of the European Union, 2019), 93, https://data.europa.eu/doi/10.2838/307887.

[655] See for example, CJEU (21 July 2011), Case C-21/10 of Károly Nagy v. Mezőgazdasági és Vidékfejlesztési Hivatal, § 47; CJEU (18 July 2013) Case C-356/12 of Wolfgang Glatzel v. Freistaat Bayern (Glatzel), § 61. Martin, 'Article 20 CFR'. As Martin observes: "Article 20 CFR is thus not a provision aiming to extend the prohibition of discrimination beyond a list of specific prohibited grounds to any ground of discrimination. As held by the Court, it simply embodies the unwritten general principle of equal treatment, which in fact constitutes a basic— and indispensable— general principle of law in any democratic state and merely reflects the Aristotelian principle that 'likes should be treated alike'." This interpretation Martin juxtaposes to equality before the law found within the U.S. Constitution, which has been interpreted as giving protection against discrimination on any ground. However, it seems that the latter interpretation is more closely linked to the goals pursued by Article 21 (1) CFR.

[656] Bell, 'Article 20: Equality Before the Law', 2014, 563.

[657] Bell, 'Article 20: Equality Before the Law', 2021, 607.

The general principle of equality tends to concern those cases that involve the distribution of scarce social, prized or other justice-relevant goods.[658] Article 20 CFR can therefore cover a variety of instances that concern the distributional choices of public authorities in areas like employment, social welfare and healthcare.[659] Yet, equality before the law not only relates to choices made by public authorities in the distribution of scarce resources. Equality before the law may also cover inequalities that follow from the operation of the law. For instance, EU criminal law is one area of importance where individuals could benefit from having the opportunity to scrutinize differential treatment: why was I made subject to a judicial process whereas others were not?[660] Hence, Article 20 CFR could become an instrument to evaluate algorithmic decision-making by public institutions in areas governed by EU law, such as data processing in law enforcement, and the criteria derived through the use of algorithms to make decisions therein.

Against this background, Article 20 CFR may act as a broad, sweeping safety net applicable to any type of differentiation, even if the acts in question appear neither morally reprehensible, nor disadvantageous to a given historically disadvantaged group.[661] Under the view set out in Section 5.1.2, Article 20 CFR would not endorse a substantive conceptualisation of equality, but rather act as procedural and democratic check which prevents the EU and national authorities from imposing differential treatment without good reason. Put differently, as long as a situation is governed by Article 20 CFR, it could become an interpretative lens for the evaluation of a differentiating act, and more specifically, whether the latter satisfies the standard test of justification or proportionality.[662] As any form of differentiation could come into conflict with the general principle of equal treatment, the main evaluative weight would then take place during the justification phase.[663]

---

[658] Bell, 'Article 20: Equality Before the Law', 2014, 563–64.

[659] Bell, 564.

[660] Bell, 564.

[661] Bell, 573.

[662] Tridimas, *The General Principles of EU Law*, 62. However, in our analysis of those reasons, equality, seen as a protective mechanism against arbitrary interference, may nonetheless be given further substance. They will however be discussed below, in the justification section. As such, Article 20 CFR represents a long tradition of EU case-law. See for example Tridimas' historical analysis of the general principle of equality. See also: Sottiaux, 'Het Gelijkheidsbeginsel : Langs Oude Paden En Nieuwe Wegen (Artikel, 2008) [WorldCat.Org]'. Tridimas, *The General Principles of EU Law*, chaps 2, The Principle of Equality.

[663] As mentioned in Section 5.1.2., whereas Article 20 CFR could function as a "safety" net, Article 21 CFR would represent a more substantive approach towards equality, where a firmer stance is taken vis-à-vis a set of clearly prohibited forms of differentiation. Bell remarks however that while this seems to be the case in practice, there is no

Article 21 CFR

Article 21 (1) CFR prohibits any type of discrimination based on any ground. While several grounds are explicitly enumerated, like Article 20 CFR - and Article 14 ECHR to which it most closely corresponds - Article 21 (1) CFR has an open-ended structure.

CLOSED INTERPRETATION –Article 21 (1) CFR is formulated as a general prohibition to discriminate. Typically, clauses that represent such a non-discrimination rationale carry within them a more substantive conceptualisation towards equality, as they aim to define more firmly forms of differentiation that are prima facie morally reprehensible.[664] If such a role is exhibited by Article 21 (1) CFR, the CJEU might be more cautious in allowing new forms of differentiation – beyond those mentioned - to be captured under its protective scope. This sentiment has already been reflected in the Court's general attitude to not expand the EU Equality Directives in covering new grounds through the Charter.[665] Ward remarks how in non-discrimination cases the CJEU will first assess whether the differentiation ground relied upon enjoys protection in either EU legislative measures or primary EU law. If such protection has not yet been offered to a ground, it will not find so through the Charter. As a consequence, "the words 'such as' in Article 21(1) of the Charter [may have been rendered devoid of any jurisprudential effect in the absence of such measures."[666] A closed reading of Article 21 CFR might also be in line with the explanations to the Charter. They clarify that the Charter should not be seen as laying down a sweeping ban of discrimination.[667] This closed-nature of Article 21 CFR somewhat contrasts with the open-ended approach found in

---

reason why Article 20 could not find application in relation to discrimination grounds mentioned under Article 21. Bell, 'Article 20: Equality Before the Law', 2014, 576–77.

[664] See also: Gerards, 'The Discrimination Grounds of Article 14 of the European Convention on Human Rights'.

[665] See Section 5.1.2 The EU Equality Directives

[666] Ward, 'The Impact of the EU Charter of Fundamental Rights on Anti-Discrimination Law', 39.

[667] The explanations also clarify the relationship between Article 19 TFEU and Article 21 CFR: "There is no contradiction or incompatibility between paragraph 1 and Article 19 of the Treaty on the Functioning of the European Union which has a different scope and purpose: Article 19 confers power on the Union to adopt legislative acts, including harmonisation of the Member States' laws and regulations, to combat certain forms of discrimination, listed exhaustively in that Article. Such legislation may cover action of Member State authorities (as well as relations between private individuals) in any area within the limits of the Union's powers. In contrast, the provision in Article 21(1) does not create any power to enact anti-discrimination laws in these areas of Member State or private action, nor does it lay down a sweeping ban of discrimination in such wide-ranging areas. Instead, it only addresses discriminations by the institutions and bodies of the Union themselves, when exercising powers conferred under the Treaties, and by Member States only when they are implementing Union law. Paragraph 1 therefore does not alter the extent of powers granted under Article 19 nor the interpretation given to that Article." Explanations relating to the Charter of Fundamental Rights (14 December 2007) (2007/c 303/02) OJ C303/17, Explanations on Article 21 – Non-discrimination.

Article 14 ECHR case-law. On that point, AG Jaaskinnen has observed how constitutional provisions, like Article 14 ECHR or Protocol 12 to the ECHR should not be used to oblige Member States to combat discrimination on grounds not spelled out in Treaties or EU legislation.[668]

OPEN INTERPRETATION - Some authors argue that the original intent behind Article 21 CFR would support the interpretation of this provision as being an open norm.[669] Eklund and Kilpatrick stress the potential of Article 21 CFR in becoming an "additional source around which to frame a rich and recursive set of EU anti-discrimination practices", and especially where grounds-based differentiation would be concerned.[670] Not only would it expand the range of the Charter as a human rights instrument, a flexible interpretation would also be consistent with the CJEU's more open-ended approach towards Article 20 CFR to which Article 21 CFR gives concrete expression.[671] Under this view, Article 21 CFR is a primary source human right that combats status-based discrimination.[672] Similarly, in light of the highly fluent nature of categorisation practices, an open-ended interpretation of Article 21 (1) CFR could play a crucial role in the evaluation of algorithmically generated harms.[673]

A GUIDING FUNCTION? - As an illustration of the guiding function Article 21 CFR can have, reference can be made to a recent ruling concerning the validity of the PNR Directive.[674] The PNR Directive governs the use of personal name records data for the prevention, detection, investigation and prosecution of terrorist offences and serious crime.[675] The Directive prohibits discriminatory profiling and more specifically, the processing of PNR data against pre-determined criteria on the basis of a person's race or ethnic origin, political opinions, religion or philosophical beliefs, trade

---

668 Opinion of Advocate General Jääskinen (17 July 2014) Case C-354/13 Kaltoft, ECLI:EU:C:2014:2106, § 24.

669 Dagmar Richter, 'Overweight and Obesity as Novel Grounds of Discrimination', in *The European Union as Protector and Promoter of Equality*, ed. Thomas Giegerich, European Union and Its Neighbours in a Globalized World (Cham: Springer International Publishing, 2020), 326, https://doi.org/10.1007/978-3-030-43764-0_16.

670 Eklund and Kilpatrick, 'Article 21 EU Charter of Fundamental Rights', 17.

671 Ward, 'The Impact of the EU Charter of Fundamental Rights on Anti-Discrimination Law', 37.

672 Eklund and Kilpatrick, 'Article 21 EU Charter of Fundamental Rights', 631–32.

673 Gerards et al., *Algorithmic Discrimination in Europe*, 141–42. Similar arguments have been made in : Gerards and Zuiderveen Borgesius, 'Protected Grounds and the System of Non-Discrimination Law in the Context of Algorithmic Decision-Making and Artificial Intelligence', 2 November 2020; Naudts, 'AI En Algoritmische Categorisatie: Gelijkheid En Non-Discriminatie'; Naudts, 'Criminal Profiling and Non-Discrimination: On Firm Grounds for the Digital Era?'

674 CJEU (21 June 2022), Case C-817/19 Ligue des Droits Humains v. Conseils des Ministres; ECLI:EU:C:2022:491.

675 Directive (EU) 2016/681 of the European Parliament and of the Council of 27 April 2016 on the use of passenger name record (PNR) data for the prevention, detection, investigation and prosecution of terrorist offences and serious crime; *OJ L 119, 4.5.2016*.

union membership, health, sexual life or sexual orientation.[676] In his opinion, AG Pitruzzella argued that the general prohibition on discriminatory profiling must be understood as pertaining to all differentiation grounds referred to in Article 21 CFR, and not only those mentioned in the Directive.[677] Unfortunately, the CJEU did not address this particular comment of the AG directly.[678] If AG Pitruzzella's reasoning would have been followed, it would not (yet) result in a fully open-ended approach towards equality. Indeed, the prohibition would remain limited to those grounds that are explicitly listed. Still, even in the case protection would be limited to a select set of status-based grounds, this evolution is to be welcomed as it could strengthen the position of EU citizens against various nefarious forms of differentiation in a wide variety of domains, including the development and deployment of algorithmically guided decision-making processes, in areas governed by EU law.

---

[676] Article 6 (4) PNR Directive.

[677] Opinion of Advocate General Pitruzella (27 january 2022) Case C-817/19 Ligue des Droits Humains v. Conseil des Ministres; ECLI:EU:C:2022:65 , § 227. For instance, the grounds "age" and "property" were not listed in Article 6 (4) PNR Directive. CJEU (21 June 2022), Case C-817/19 Ligue des Droits Humains v. Conseils des Ministres; ECLI:EU:C:2022:491., § 197 – 199. In the case of G.D., the CJEU had to consider the use of geographic location as a criterion for data retention in the fight against serious crime: "a targeted measure for the retention of traffic and location data may, at the choice of the national legislature and in strict compliance with the principle of proportionality, also be set using a geographical criterion where the competent national authorities consider, on the basis of objective and non-discriminatory factors, that there exists, in one or more geographical areas, a situation characterised by a high risk of preparation for or commission of serious criminal offences. Those areas may include places with a high incidence of serious crime, places that are particularly vulnerable to serious crime, such as places or infrastructure which regularly receive a very high volume of visitors, or strategic locations, such as airports, stations, maritime ports or tollbooth areas." The Court added that: "competent national authorities may adopt […] a targeted measure of retention using a geographic criterion, such as, inter alia, the average crime rate in a geographical area, without that authority necessarily having specific indications as to the preparation or commission, in the areas concerned, of acts of serious crime." According to the CJEU "a targeted retention using that criterion is likely to concern […] both the areas marked by a high incidence of serious crime and areas particularly vulnerable to the commission of those acts." The use of geographic location the CJEU therefore believes, is unlikely to "give rise to discrimination, as the criterion drawn from the average rate of serious crime is entirely unconnected with any potentially discriminatory factors." CJEU (5 April 2022) Case C-140/20 G.D. v. Commissioner of An Garda Siochana; ECLI:EU:C:2022:258, § 79-80

[678] In their ruling, the CJEU did not clearly affirm this "expansion" of grounds listed in the PNR Directive to match those of Article 21 CFR, the CJEU did observe that the criteria used during profiling should 'in no circumstances' be based on the ones listed in the PNR Directive. The Court did observe that the PNR Directive should be read in light of Article 21 CFR, and that the Charter prohibits both direct and indirect discrimination.Hence, any criteria used, even those worded neutrally, should not place persons having those protected characteristics at a particular disadvantage.

5.3 The Justification of Differentiation

In Chapter 3, I argued how, next to the law's initial range of protection, the justification phase co-shapes the law's conceptualisation of equality.[679] In order to uncover the structuring function equality holds within a legal order - and via the law, within society at large - we should gain insight into the review mechanisms that are applied when equality is positioned against competing interests. I noted how this investigation should be further specified in order to successfully assess the law's ability to address the egalitarian harms algorithms risk to impose. First, I look into the structure of the review process: are there different justification procedures, and if so, does their evaluative focal point differ? Second, I want to discover the particular values equality represents when it is positioned against competing interests: what (normative) rationales explain the legal and social function equality performs? Through those findings I can establish whether the law can address the moral, socio-relational, and distributive interests of decision-subjects.

As a general principle of EU law, the position held by the principle of equal treatment is characterised by great complexity. Because the regulatory landscape is as scattered, it is difficult to source reliable and structured guidance on the EU's equality review. Though I cannot disentangle this fragmentation within the confines of this dissertation, in the following Sections, I aim to facilitate the (normative) navigation of the EU's equality review by shedding light on the aforementioned elements of the justification phase. I first address the EU Equality Directives, which give concrete expression to the general principle of equality. I focus on the egalitarian objectives these instruments sought to achieve. In addition, I also analyse the justification mechanisms they put forward. I then turn to the structural components that shape the equality review traditionally applied by the CJEU in their interpretation of the general principle of equal treatment. In the final Section, I examine how the CJEU, with the support of its Advocate Generals, has given additional and overarching (normative) substance to the concept of equality in its interpretation and analysis of the EU Equality Directives and the Charter. Next to the normative objectives pursued by EU lawmakers, I assess whether the rationales relied upon by the CJEU to further ground the general principle of equal treatment either enhance or obstruct its capacity to evolve and transform in light of social (and technological) change. During my analysis, I concurrently investigate to what extent choices made by either lawmakers or the judiciary affect the law's capacity to address algorithmic harms.

[679] See Chapter 3

5.3.1 Justification in the EU Equality Directives

The EU Equality Directives have further concretised the function of equality within the Union. In complex situations, the general principle of equal treatment is not always capable of demarcating the dividing line between those acts of differentiation that are justifiable and those that are not. It might be appropriate to weed out "straightforward, crude or arbitrary forms of discrimination", but combatting discrimination becomes most effective once appropriate definitions have been elaborated upon.[680] The EU Equality Directives, which form a specific expression to Article 21 (1) CFR, define, and lay down the rules, designed to eliminate discrimination, and in doing so, they specify what society accepts as justifiable differentiation: they "classify the application of a particular criteria [differentiation ground] in particular circumstances as acceptable or unacceptable and […] give binding legal effect to that classification."[681] The Directives can therefore be considered the (most) concrete (albeit not complete) manifestation of equality as a substantive fundamental and human right.[682]

Although the CJEU views the EU Equality Directives as the expression of the fundamental right of equality enshrined within the Charter, they are still the result of a long and phased political decision-making process.[683] When, and under what conditions, inequalities are justified should therefore be initially sought in the motivations of the EU legislator.[684] Upon their introduction, the EU Equality Directives initiated a dynamic shift in relation to the position and function equality

---

[680] Opinion of AG Sharpston (22 May 2008) Case-427/06 Birgit Bartsch v Bosch und Siemens Hausgeräte (BSH) Altersfürsorge GmbH; ECLI:EU:C:2008:297; § 49.

[681] Opinion of AG Sharpston (22 May 2008) Case-427/06 Birgit Bartsch v Bosch und Siemens Hausgeräte (BSH) Altersfürsorge GmbH; ECLI:EU:C:2008:297; § 49 and 56.

[682] While the EU Equality Directives are the clearest expression of unjustified differentiation, the general principle of equality does a lot of preparatory groundwork. This relationship between the general principle of equality and the adoption of specific non-discrimination laws was clarified by AG Sharpston. The need for an EU Equality Directive emerges when a new possible scope of the principle is discovered. Such she believed was the case for the recognition of age discrimination as an unjustified form of discrimination: "It is precisely because the general principle of equality has now been recognised also to include equality of treatment irrespective of age that an enabling legislative provision such as [Article 19 TFEU] becomes necessary and is duly used as the basis for detailed legislative intervention." § 49-50.

[683] Xenidis, 'Transforming EU Equality Law?' Belavusau and Henrard, 'A Bird's Eye View on EU Anti-Discrimination Law'. Since their arrival at the beginning of the 2000's, no new concrete Equality Directives have been adopted or proposed, which has been attributed to a changed political landscape and a general reluctance among too many Member States to expand upon the protection of EU Equality law.

[684] See also observations made under 5.2.1, and the observations of AG Geelhoed in the case of *Chacón Navas* in particular. Opinion of AG Geelhoed (16 March 2006) in Case C-13/06 Sonia Chacón Navas v. Eurest Colectividades SA *(Chacón Navas).*

had held within the EU; equality as a tool for market integration was suddenly complemented with a more socially-oriented dimension of citizenship through identity-based non-discrimination rights.[685] The EU Equality Directives were therefore crucial in equality's evolution from a market-oriented, distributional and formal principle, to one that is more substantive in nature, highlighting the complementary relationship between equality and human dignity.[686] They remove illegitimate barriers to equal opportunity in order to ensure individuals' participation to essential social institutions or publicly prized goods, such as a person's access to the labour market or access to goods and services.[687] Drawing on Xenidis, the EU Equality Directives can be seen as a site of reconciliation, whereby equal citizenship is realised through and before the market.[688] In equality *through* the market, the EU system facilitates individual economic and social participation through distributive safeguards aimed at correcting unbalanced market opportunities.[689] Equality *before* the market ensures the equal recognition of vulnerable and marginalised groups by granting identity-based (read grounds-based) non-discrimination rights.[690] Both narratives are interwoven within the Directives: "While the substance of non-discrimination rights is made of distributive opportunities, the recognition paradigm informs the grammar of their allocation."[691] Socio-relational equality is realised within the confines of and through clearly delineated distributive objectives.[692]

---

[685] See for instance: Xenidis, 'Transforming EU Equality Law?', 56; Bell, 'The Transformation of EU Anti-Discrimination Law', 196.

[686] Belavusau and Henrard, 'The Impact of the 2000 Equality Directives on EU Anti-Discrimination Law : Achievements and Pitfalls', 9.

[687] Raphaële Xenidis, 'The Polysemy of Anti-Discrimination Law: The Interpretation Architecture of the Framework Employment Directive at the Court of Justice.', *Common Market Law Review* 58 (2021): 1656.

[688] Xenidis, 'Shaking the Normative Foundations of EU Equality Law', 19.

[689] Xenidis, 21. See for example also: AG Kokott's opinion in Achbita: "On the one hand, the principle of equal treatment enshrined in Article 21 of the Charter of Fundamental Rights and given specific expression in Directive 2000/78 is expressly intended to make it easier for disadvantaged groups to gain access to employment and occupation, in order thus to promote their participation in economic, cultural and social life and the realisation of their potential." Opinion of AG Kokott (31 May 2016) Case C-157/15 Samira Achbita and Centrum voor Gelijkheid van Kansen en voor Racismebestrijding v. G4S Secure Solutions NV. , § 123.

[690] Xenidis, 21.

[691] Xenidis, 'Transforming EU Equality Law?', 29.

[692] Some have even argued that under this market focus equality is held hostage by a neoliberal agenda. See: Alexander Somek, 'Moral Re-Engineering: The Neoliberal Predicament', in *Engineering Equality* (Oxford: Oxford University Press, 2011), https://doi.org/10.1093/acprof:oso/9780199693375.003.0008. In the case of race and gender inequality, the Directives do push towards a more comprehensive and collective ideal of justice and equality, and market ideals are not as dominant. Xenidis, 'Shaking the Normative Foundations of EU Equality Law', 31. See for example: Opinion of AG Kokott (31 May 2016) Case C-157/15 Samira Achbita and Centrum voor Gelijkheid van Kansen en voor Racismebestrijding v. G4S Secure Solutions NV. , § 122-123.

While social and cultural diversity, inclusion and recognition inform the attribution of non-discrimination rights, the boundaries of the Directives' equality concept are tie to and constrained by the law's distributive dimension. These limitations regarding the Directives' reach also affect their potential in governing algorithmic egalitarian harms. When discriminatory algorithms interfere with the desired distribution of social and economic opportunities as defined by the EU Equality Directives, they are indeed likely to produce a socio-relational egalitarian harm. When people are barred from those aspects of life that have been identified as critical for social and economic participation, it undermines their status as social equals. Yet, under the Directives' distributive structure, these socio-relational harms only find protection when they can be tied to an allocative harm. It is highly doubtful however whether socio-relational ideals, such as the respect for cultural and social diversity, can be realised by regulating how people should treat one another – *in casu*, a prohibition to treat others as lesser human beings for possessing a certain characteristic - in relation to a desired state of distributive affairs alone. Put differently, the socio -relational dimension has been given little independence. Hence, the Directives' approach may fail to recognise algorithmic socio-relational inequality as a stand-alone harm that may arise both within and outside the domains captured under their material scope.[693]

Next to these normative underpinnings, the justification of (in)equalities is further defined by the EU Equality Directive's distinction between direct and indirect discrimination.[694]

---

[693] For more information on socio-relational algorithmic harms, see also Chapter 6: "Socio-Relational Conceptualisations of Equality and Algorithmically Guided Decisions

[694] Much has been said already about this distinction and whether or not it should be maintained. Scholars have observed how the boundaries between both notions, both in theory and practice, are blurred, rather than clearly defined, and that perhaps one should move beyond this dichotomy. Sandra Fredman, 'Direct and Indirect Discrimination: Is There Still a Divide?', in *Foundations of Indirect Discrimination Law*, ed. Hugh Collins and Tarunabh Khaitan, 1st ed. (Oxford: Hart Publishing, 2018), 31–56, http://www.bloomsburycollections.com/book/foundations-of-indirect-discrimination-law/ch2-direct-and-indirect-discrimination-is-there-still-a-divide/. De Vos, 'The European Court of Justice and the March towards Substantive Equality in European Union Anti-Discrimination Law'. Fredman, 'Direct and Indirect Discrimination: Is There Still a Divide?'; Hugh Collins and Tarunabh Khaitan, 'Indirect Discrimination Law: Controversies and Critical Questions', in *Foundations of Indirect Discrimination Law*, ed. Hugh Collins and Tarunabh Khaitan, 1st ed. (Oxford: Hart Publishing, 2018), 1–30, http://www.bloomsburycollections.com/book/foundations-of-indirect-discrimination-law/ch1-indirect-discrimination-law-controversies-and-critical-questions/; Nicholas Bamforth, 'Approaching the Indirect–Direct Discrimination Distinction: Concepts, Justifications and Policies', in *Foundations of Indirect Discrimination Law*, ed. Hugh Collins and Tarunabh Khaitan, 1st ed. (Oxford: Hart Publishing, 2018), 57–82, http://www.bloomsburycollections.com/book/foundations-of-indirect-discrimination-law/ch3-approaching-the-indirect-direct-discrimination-distinction-concepts-justifications-and-policies/. In addition to other contributions

The Justification of Direct Discrimination

The EU Equality Directives define direct discrimination as the situation where a person is treated less favourably than another is, has been, or would be treated in a comparable situation, on any of the characteristics listed within the Directives. For instance, the Austrian employment algorithm, which systematically assigned a lower score to women, women with children, disabled persons, and people over thirty, constituted a case of direct algorithmic discrimination. Because of their lower score, these persons were less likely to receive adequate welfare support.[695] As a standard rule, acts categorised as directly discriminatory are prohibited and cannot be justified. Only where the Directives foresee in an express exemption could direct discrimination be allowed. These situations are quite scarce and phrased in a narrow sense.[696]

The Directives' definition of direct discrimination exemplifies the formal approach towards equality. It requires a comparator group to be present, as well as a causal link between the

found within: Hugh Collins and Tarunabh Khaitan, eds., *Foundations of Indirect Discrimination Law*, 1st ed. (Oxford: Hart Publishing, 2018), https://doi.org/10.5040/9781509912551. AGs too can differ in their opinion as to when a specific act should be considered directly or indirectly discriminatory. Advocate Generals Sharpston and Kokott differed in opinion on how to classify a workplace regulation that prohibits employees from wearing religious signs or apparel. Opinion of AG Sharpston (13 July 2016) Case C-188/15 Asma Bougnaoui and ADDH v Micropole SA; Opinion of AG Kokott (31 May 2016) Case C-157/15 Samira Achbita and Centrum voor Gelijkheid van Kansen en voor Racismebestrijding v. G4S Secure Solutions NV. See also: Matthias Mahlmann, 'ECJ Headscarf Series (3): The Everyday Troubles of Pluralism', *Strasbourg Observers* (blog), 12 September 2016, https://strasbourg.weichie.dev/2016/09/12/ecj-headscarf-series-3-the-everyday-troubles-of-pluralism/.

Furthermore, and as highlighted, In discussions concerning the role of EU Equality law as a means to govern algorithmically profiling practices, the distinction also appears more difficult to maintain, For a more concrete analysis as to whether the direct/indirect distinction makes for a good conceptual fit within the algorithmic environment, see also the work of Gerards et al., *Algorithmic Discrimination in Europe*; Xenidis and Senden, 'EU Non-Discrimination Law in the Era of Artificial Intelligence : Mapping the Challenges of Algorithmic Discrimination'.

[695] Lopez, 'Bias Does Not Equal Bias'.

[696] For example, and with the exception of the Gender Goods and Services Directive, all EU Equality Directives allow discrimination on the basis of a protected characteristic where it this can be used as a "genuine occupational requirement." Moreover, even in the case an exception applies, it is upon the condition that the deviating measure must be legitimate and proportionate. This also allows an employer to discriminate on the basis of a protected trait where there is "an inherent link with the capacity to perform or the qualifications required for a particular job." See: Article 4 (1), Framework Employment Directive. Article 14 (2) Gender Recast Directive; Article 4 Racial Equality Directive. There are two other exceptions possible. First, discrimination on the basis of religion or religious beliefs may be possible by "faith-based" organisations. Second, the Framework Employment Directives allows certain forms of age discrimination (Article 6 Framework Employment Directive). See also Europäische Union, Europarat, and Europäischer Gerichtshof für Menschenrechte, *Handbook on European Non-Discrimination Law*, 96.

differentiating act and the protected characteristic.[697] This definition primarily reflects the idea that the characteristics in question should be considered irrelevant for decision-making: people should not be treated unfavourably on the basis of the protected characteristics in the areas governed by the Directives. For example, in decisions concerning access to employment and occupation, no disadvantage should be imposed onto people due to their religion or beliefs. Decision-makers should try to make abstraction of these characteristics.[698] While the EU Equality Directives look to mitigate the unfair imposition of socio-economic disadvantage through the principle of equal treatment, the wrongful nature of discrimination, and the unjustified nature of the inequalities in question, lies in the *use* of those specific characteristics.[699] Though examples of directly discriminatory algorithms can be found, it is believed that the concept may lose out on relevance.[700] For one, due to their awareness of their legal obligations, decision-makers are unlikely to make direct use of protected grounds as a decisive element in decision-making processes.[701] In addition, algorithmic differentiation grounds are often complex and non-representational. Even in the case these non-representational grounds act as a proxy for a protected trait, their use is more likely to be categorised as a form of indirect discrimination (see below).

Be that is it may, the formal direct discrimination approach has, in the words of De Vos, been "bent" towards more substantive, socio-relational aims.[702] For instance, the CJEU has diluted the victim-comparator-causation trifecta associated with formal equality to give direct discrimination a more socio-relational push. Consequently, certain cases of differential treatment that would have found justification otherwise, were re-categorised as a direct, and consequently unjustifiable, form of discrimination.[703] This enables the CJEU to re-categorise seemingly neutral

---

[697] See also discussions above regarding the problematic nature of the comparator approach for evaluating algorithmic decision-making "Chapter 4, Section 4.3.1 The Review Procedure". Fredman, 'Substantive Equality Revisited', 717–19. De Vos, 'The European Court of Justice and the March towards Substantive Equality in European Union Anti-Discrimination Law', 66.

[698] Fredman, 'Substantive Equality Revisited', 732–33.

[699] Again, where racial equality is concerned, market ideals appear less dominant, see footnote:

[700] Xenidis and Senden, 'EU Non-Discrimination Law in the Era of Artificial Intelligence : Mapping the Challenges of Algorithmic Discrimination'; Gerards et al., *Algorithmic Discrimination in Europe*, 69.

[701] Of course, in certain cases, the use of such parameters might be legitimate, e.g., one could make use of protected characteristics to detect whether a decision-making system does not discriminate by proxy.

[702] De Vos, 'The European Court of Justice and the March towards Substantive Equality in European Union Anti-Discrimination Law', 67.

[703] See also Section 5.2.1 "Extending the Meaning of Differentiation Grounds."

differentiation grounds that, under a veil neutrality, were used to cover direct discrimination.[704] In similar fashion, CJEU could loosen the causality requirement to better capture algorithms' correlation-based reasoning.[705] Examples of such developments have already been examined: the CJEU may declare the use of a particular characteristic prohibited when it is inextricably linked to a protected characteristic[706], it may extend the Directives' personal scope via the discrimination by association doctrine[707], it may eliminate the need for a comparator[708], etc.

The Justification of Indirect Discrimination

The EU Equality Directives define indirect discrimination as the situation where an apparently neutral provision, criterion or practice places a protected person or group at a particular disadvantage when compared to others. Indirect discrimination is often described as forming a departure from the formal approach to one that is more substantive. It focuses on the *impact or effect* a seemingly neutral rule has, rather than the direct treatment such a rule imposes. As society tends to be structured by, and to the benefit of, dominant social groups, neutral measures are frequently disadvantageous to those that are underrepresented. It is therefore said that the indirect discrimination doctrine reflects a social engineering rationale: it requires decision-makers to consider the actual (social) context or environment in which a neutral rule will operate.[709]

The notion of indirect discrimination is often seen as a promising concept for the regulation of algorithmic discrimination. Algorithmic differentiation grounds are often defined as neutral rules

---

[704] Xenidis and Senden, 'EU Non-Discrimination Law in the Era of Artificial Intelligence : Mapping the Challenges of Algorithmic Discrimination', 19.

[705] See for instance: Raphaële Xenidis, 'Tuning EU Equality Law to Algorithmic Discrimination: Three Pathways to Resilience', *Maastricht Journal of European and Comparative Law* 27, no. 6 (December 2020): 747, https://doi.org/10.1177/1023263X20982173.

[706] See for example, CJEU (26 January 2021) Case C-16/19 VL v. Szpital Kliniczny im. dra J. Babińskiego Samodzielny Publiczny Zakład Opieki Zdrowotnej w Krakowie.

[707] See also Section 5.2.1 "Extending the Meaning of Differentiation Grounds.", and the discussion of the cases of *Coleman* and *Chez.*

[708] See for instance: CJEU (8 November 1990) Case C-177/88 Dekker, ECLI:EU:C:1990:383.

[709] Dagmar Schiek, 'Indirect Discrimination', in *Materials, Cases and Text on National, Supranational and International Non-Discrimination Law*, ed. Dagmar Schiek, Lisa Waddington, and Mark Bell, IUS COMMUNE Casebooks on the Common Law of Europe (Oxford: Hart Publishing, 2007), 330. De Vos notes how the phrasing of the Directives, which requires decision-makers to assess whether a neutral rule *would* place people at a disadvantage, holds parties liable for the sheer social reality they find themselves in, regardless of their own role or contribution to this reality. De Vos, 'The European Court of Justice and the March towards Substantive Equality in European Union Anti-Discrimination Law', 70.

(or grounds). Still, these grounds may serve as a proxy for a protected trait, and when they do, they are often most detrimental to historically disadvantaged or marginalised communities. Hence, the indirect discrimination doctrine may be used to tackle the problem of algorithmic proxies in cases where they disadvantage people on the basis of protected characteristics.[710] Unlike direct discrimination however, indirect discrimination can be objectively justified. This justification possibility is entirely open-ended and protection against indirect algorithmic discrimination depends entirely on the CJEU's equality review. More specifically, indirect discrimination can be justified by demonstrating that the provision, criterion, or practice in question pursued a legitimate aim, and the means used to pursue those aims were proportionate and necessary.[711] As this justification assessment resembles a traditional proportionality approach, which is applied in relation to both the Directives and the Charter, it will be discussed below.

5.3.2 The EU's Equality Review

In their interpretation of existing and upcoming EU legislation governing algorithmically guided decision-making systems, including the EU Equality Directives, the CJEU can further clarify the

[710] Still, the application of the indirect discrimination doctrine is not without problems. Difficulties may for example arise in the definition of the disadvantaged group, the disadvantage incurred by that group, and the demonstration of said disadvantage, including the means through which it can be demonstrated. For an exploration of these topics, in relation to algorithms, reference can be made to the work of Wachter, and more particularly: Wachter, Mittelstadt, and Russell, 'Why Fairness Cannot Be Automated'.

[711] See for example Article 2 § 2. Council Directive 2000/78/EC of 27 November 2000 establishing a general framework for equal treatment in employment and occupation; Available via: https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=celex%3A32000L0078. Rather than a proportionality approach, some argue in favour of an approach where indirect discrimination is established through a causal link test. The objective justification could then be seen as rebutting the presumption that a causal link exists between a neutral rule or practice or a mere structural impact. Schiek, 'Indirect Discrimination', 439. For a more thorough assessment of the objective justification test, including the distinction between "justification proper" and "causation", see also: Christa Tobler, *Indirect Discrimination: A Case Study into the Development of the Legal Concept of Indirect Discrimination under EC Law*, Social Europe Series, v. 10 (Antwerpen: Intersentia, 2005). Schiek has argued that the possibility to justify indirect discrimination under EU law was a precondition for having the indirect discrimination prohibition applied onto private actors. This does allow decision-makers to impose an unequal distribution of resources and opportunities. In so doing, the free market is giving priority in its ability to plan adequate distributions. Schiek, 'Indirect Discrimination', 330. See also: Christa Tobler, 'Limits and Potential of the Concept of Indirect Discrimination' (European Network of Legal Experts in the non-discrimination field, 2008), 32–33.

conditions under which algorithmic differentiation acts are unjust.[712] In this Section, I aim to clarify to what extent structural choices made during the CJEU's equality review can affect the law's ability to address algorithmic harms – whether the CJEU decides to expand the Charter's scope to cover non-protected grounds or not.

The Standard Formula

The CJEU has built a long tradition of case-law treating the general principle of equality as a principle structured around a balancing of interests.[713] Under its standard formula, the general principle of equal treatment, as enshrined within Article 20 and 21 CFR, and given further expression through the Directives, requires that comparable situations must not be treated differently and that different situations must not be treated in the same way unless such treatment is objectively justified.[714] The wording and phrasing of the CJEU, and notions involved (i.e.,

[712] In this regard, several cases are notable. For instance, in the Leger case, which did not concern the application of one of the Equality Directives, the CJEU did establish that in their implementation of EU law, Member States should guarantee the principle of equality, and more specifically, protect individuals against being discriminated on one of the grounds listed in Article 21 (1) CFR. The case concerned a Member State's decree which prohibited men who had engaged in sexual relationships with other men from donating blood. The CJEU found that such a decree may discriminate against gay people on grounds of sexual orientation within the meaning of Article 21 (1) CFR. CJEU (29 April 2015) Case C-528/13 Léger. A similar watershed moment concerned Egenberger case, where the CJEU confirmed the direct horizontal effect of Article 21 CFR. The CJEU first observed that: "The prohibition of all discrimination on grounds of religion or belief is mandatory as a general principle of EU law. That prohibition, which is laid down in Article 21(1) of the Charter, is sufficient in itself to confer on individuals a right which they may rely on as such in disputes between them in a field covered by EU law." And second, it held that: "As regards its mandatory effect, Article 21 of the Charter is no different, in principle, from the various provisions of the founding Treaties prohibiting discrimination on various grounds, even where the discrimination derives from contracts between individuals." CJEU (17 April 2018) Case C-414/16 Egenberger v. Evangelisches Werk Fur Diakonie und Entwicklung eV; ECLI:EU:C:2018:257, § 76 and 77. In the Case of Test-Achats, the CJEU declared incompatible a Directive provision which allowed Member States to exempt ensures from having to provide unisex premiums and benefits to men and women. This clause went against the goal of equal treatment between men and women and was therefore incompatible Article 21 and 23 CFR. CJEU (1 March 2011) Case C-236/09 Test-Achats; ECLI:EU:C:2011:100.

[713] As a reminder, the ECtHR analyses equality cases through a multi-pronged approach. First, in order for an issue to arise, there must either be a difference in treatment of persons in analogous or relevantly similar situations (or vice versa). Second, the ECtHR would analyse whether this differential treatment could objectively justify. Here too, there is a structure of comparability and objective justification. See also: Opinion of AG Sharpston (13 July 2016) Case C-188/15 Asma Bougnaoui and ADDH v Micropole SA, § 121. Muir, 'The Essence of the Fundamental Right to Equal Treatment', 822.

[714] For instance, in the case of Milkova, the CJEU clarified how "the principle of equal treatment is a general principle of EU law, now enshrined in Articles 20 and 21 of the Charter, which requires that comparable situations must not be treated differently and that different situations must not be treated in the same way unless such treatment is objectively

comparability, legitimate aim, appropriateness and necessity), when interpreting the Charter's equality provisions, (roughly) correspond to the legal justification requirements imposed by the EU Equality Directives onto cases of indirect discrimination. Insofar as both review procedures are aligned, below-mentioned observations regarding its performance for the evaluation of algorithmically generated inequalities, also apply to the CJEU's review of indirect discrimination.[715]

The CJEU has further clarified how the comparability requirement must be assessed in light of all factors that characterise the situation.[716] This assessment must not be carried out in a global and abstract manner, but in a specific and concrete manner, in light of the subject-matter and the aim of the measure creating the distinction at issue. Second, differential treatment can be objectively justified when it is based upon an objective and reasonable criterion, that is, if the difference relates to a legally permitted aim and is proportionate.[717]

---

justified." CJEU (9 March 2017) Case C-406/15 Milkova; ECLI:EU:C:2017:198, § 55. See also: CJEU (22 May 2014), Case C-356/12 *Glatzel*, § 42, 43 and 81; CJEU (17 October 2013) Case C-101/12 Herbert Schaible v. Land Baden-Wurttemberg ECLI:EU:C:2013:661, § 76-77. See also: Tridimas, *The General Principles of EU Law*; Aharon Barak, ed., 'Proportionality Stricto Sensu (Balancing)', in *Proportionality: Constitutional Rights and Their Limitations*, Cambridge Studies in Constitutional Law (Cambridge: Cambridge University Press, 2012), 340–70, https://doi.org/10.1017/CBO9781139035293.016; Muir, 'The Essence of the Fundamental Right to Equal Treatment'.

[715] See for instance also the Case of Chez, which concerned a case of indirect discrimination on the basis of ethnicity. CJEU (16 July 2015) Case C-83/14 Chez, ECLI:EU:C:2015:480, § 112, in reference to § 88. In it, the CJEU referred to the Arcelor case in order to explain how a breach of the principle of equal treatment must be assessed in light of all elements that characterise them. This case did not involve a « protected characteristics » but assessed a difference in treatment in relation to the greenhouse gas emission allowance trading scheme between the steel sector on the one hand, and chemical and non-ferrous metal sectors on the other hand. See : Case C-127/07 Société Arcelor Atlantique et Lorraine and Others v Premier ministre, Ministre de l'Écologie et du Développement durable and Ministre de l'Économie, des Finances et de l'Industrie ; ECLI:EU:C:2008:728, § 25.

[716] CJEU (9 March 2017) Case C-406/15 Milkova; ECLI:EU:C:2017:198, § 56-57. In Milkova, the CJEU found that civil servants with a disability were in a comparable situation as employees with a disability. For a case where non-comparable situations were treated similarly, reference can be made to the Zoi Chatzi case. Here, the CJEU found that parents of twins are in a special situation: the raising of two children in parallel is not comparable to the task of bringing up children of different ages. Accordingly, the special situation of twins had to be taken into account by the Member States. CJEU (16 September 2010) Case C-149/10 Zoi Chatzi v Ypourgos Oikonomikon, ECLI:EU:C:2010:534, § 67-71. CJEU (16 December 2008). See also : Case C-127/07 Société Arcelor Atlantique et Lorraine and Others v Premier ministre, Ministre de l'Écologie et du Développement durable and Ministre de l'Économie, des Finances et de l'Industrie ; ECLI:EU:C:2008:728, § 25.

[717] CJEU (9 March 2017) Case C-406/15 Milkova; ECLI:EU:C:2017:198, § 55. CJEU (5 May 2014) Case C-135/13 Szatmári Malom, EU:C:2014:327; § 67.

COMPARABILITY – In Chapter 4, I have argued how a comparability requirement - whether it is performed in the interpretation of the Charter or the Equality Directives - faces difficulty when applied onto the algorithmic environment.[718] For one, in cases of complex forms of (status-based) discrimination, such as intersectional discrimination, people might be as uniquely situated, that there is no group they can be adequately compared with. Second, the function of algorithms is precisely to group together individuals based upon (alleged) relevant similarities or differences. In this context, the argument could always be made that the groups subject to differential treatment, due to the differences that have been detected, are never in a comparable situation, and hence can be justifiably differentiated. If this argument is pushed to its limits, it could potentially obstruct any further review into the actual disadvantage carried by persons subject to differential treatment.

LEGITIMATE AIM - The notion of legitimate aim is open-ended. As a result, a broad range of acceptable justification grounds are typically available to decision-makers.[719] Little guidance is provided as to what can or cannot be considered a legitimate aim. In cases concerning indirect sex discrimination, the Court has specified that the aim pursued should not be discriminatory in itself: the aim must be based upon objectively justified factors unrelated to any discrimination on grounds of sex.[720] In cases where a disparate impact measure was imposed by an economic actor, the CJEU has further clarified that the aim pursued must correspond to a real need, for which the undertaking in question must be able to demonstrate that they themselves would face adverse consequences if the contested measure would not be introduced.[721] Regarding the legitimacy of the objectives pursued, the Charter specifies that limitations to the general principle of equality

---

[718]See observations Chapter 4, Section 4.3.1 "Comparability."

[719] Tobler, 'Limits and Potential of the Concept of Indirect Discrimination', 32–33.

[720] Likewise, in the case of Feryn – though this concerned a case of direct discrimination captured under Directive 2000/43/EC - the (alleged) discriminatory wish of customers to not be aided by people with a foreign background was not considered as an acceptable justification for a discriminatory recruitment policy excluding people on the basis of their racial and ethnic origin. CJEU (13 May 1986) Case C-170/84 of Bilka - Kaufhaus GmbH v Karin Weber von Hartz CJEU (9 February 1999), § 31 and Case C-167/97 Seymour-Smith, § 72. Tobler, 33. Schiek, 'Indirect Discrimination', 444. CJEU (10 July 2008) Case C-54/07 Centrum voor Gelijkheid van kansen en voor racismebestrijding v. Firma Feryn NV. Erica Howard, 'Headscarves and the CJEU: Protecting Fundamental Rights and Pandering to Prejudice: The CJEU Does Both', *Maastricht Journal of European and Comparative Law*, 28 February 2022, 1023263X221080557, https://doi.org/10.1177/1023263X221080557.

[721] CJEU (13 May 1986) Case C-170/84 of Bilka - Kaufhaus GmbH v Karin Weber von Hartz, § 37. CJEU (15 July 2021) Joined Cases C-804/18 and C-341/19 IX v. WABE EV and MH Müller Handels GmbH v. MJ; § 76; CJEU (26 June 2001) Case C-381/99, Susanna Brunnhofer v. Bank der österreichischen Postsparkasse AG,§ 67. CJEU (10 July 2008) CJEU (15 July 2021) Joined Cases C-804/18 and C-341/19 IX v. WABE EV and MH Müller Handels GmbH v. MJ; § 67-70.

must be provided for by law, genuinely meet objectives of general interest recognised by the Union or the need to protect the rights and freedoms of others.[722]

PROPORTIONALITY - The CJEU's core equality review concerns the proportionality of a given measure. A measure will be proportionate when it does not exceed the limits of what is appropriate and necessary.[723] Though explicated as consisting of two components, the proportionality assessment is typically envisaged as a three-step process. Under a first step, the CJEU will consider whether a measure is suitable or reasonably likely to achieve the legitimate aim (appropriateness). For instance, the CJEU has clarified that a measure is appropriate when it is actually designed to meet the objective or legitimate aim pursued.[724] In addition, the CJEU has specified how a measure must "genuinely reflect a concern to attain that aim" and "be pursued in a consistent and systematic manner."[725] Under a second step, the CJEU will assess whether there are other less restrictive measures available that have the same result (necessity). As far as necessity is concerned, the CJEU has argued that a measure should not go beyond what is necessary to achieve the aim pursued.[726] Whether a decision-maker has remained within those limitations can be assessed in light of the context and the adverse effects the measure is liable to cause for the persons concerned[727] or the goals and adverse consequences the decision-maker wishes to avoid by choosing for a given rule or practice.[728] One such contextual element may relate to whether measures have been put in place that are capable to prevent or compensate the

[722] Art. 52 (1) CFR.

[723] Tridimas, *The General Principles of EU Law*, 139. See for instance on the relationship between the principle of proportionality and non-discrimination, and the requirements of appropriateness and necessity: CJEU (23 March 2006) C-535/03 Unitymark and North Sea Fishermen's Organisation, § 56.

[724] See for instance CJEU (18 January 2018) Case C-270/16 of Carlos Enrique Ruiz Conejero v. Ferroser Servicios Auxiliaires SA, Ministerio Fiscal, § 44.

[725] See for instance CJEU (21 January 2021) Case-843/19 of INSS v. BT; §32, CJEU (12 January 2010), Case C-341/08 Petersen, § 72. See also: CJEU (5 July 2017) Case C-190/16 Werner Fries v. Lufthansa CityLine GmbH, ECLI:EU:C:2017:513, § 48.

[726] See for instance CJEU (18 January 2018) Case C-270/16 of Carlos Enrique Ruiz Conejero v. Ferroser Servicios Auxiliaires SA, Ministerio Fiscal, § 45; CJEU (23 March 2006) C-535/03 Unitymark and North Sea Fishermen's Organisation, § 56.

[727] See for instance CJEU (18 January 2018) Case C-270/16 of Carlos Enrique Ruiz Conejero v. Ferroser Servicios Auxiliaires SA, Ministerio Fiscal, § 49.

[728] CJEU (15 July 2021) Joined Cases C-804/18 and C-341/19 IX v. WABE EV and MH Müller Handels GmbH v. MJ; § 70. Some authors have argued for a strict necessity test: that is, there should be no other, less restrictive alternatives available to the decision-maker. Tobler, *Indirect Discrimination*, 241.

disadvantages affected by a given rule or practice.[729] Finally, even if it appears no less restrictive means are available, the CJEU may still verify whether the measure in question does not have an excessive effect on the interests of those affected (proportionality *stricto sensu* or the "fair balancing test").[730] When analysing the impact of a neutrality rule in the workplace on people's capacity to express their religious beliefs, AG Kokott observed: "measures must not, even if they are appropriate and necessary for achieving legitimate objective, give rise to any disadvantages which are disproportionate to the objectives pursued [and] a fair balance must be struck between the conflicting interests of employees […] on the one hand, and undertakings […] on the other."[731]

According to Tridimas, the distinction between these steps, and step two and three in particular, is not always visible, nor does the CJEU always look for less restrictive alternatives.[732] The CJEU case-law does not position the proportionality criterion as a uniform test, but, as he puts it, as a "flexible principle which is used in different contexts to protect different interests and entails varying degrees of judicial scrutiny."[733] In a similar vein, the application of the principle of equality is one of striking a balance between conflicting interests, rather than mechanical: the Court may consider the factual and legislative context at hand, the objectives of the legislation involved, the interests represented, as well as economic or social considerations at play.[734]

---

[729] See for instance CJEU (18 January 2018) Case C-270/16 of Carlos Enrique Ruiz Conejero v. Ferroser Servicios Auxiliaires SA, Ministerio Fiscal, § 55. In this particular case, the disadvantaged group were people with a disability.

[730] Tridimas, *The General Principles of EU Law*, 139. According to Fredman, such a test requires several elements of consideration. First, what is the closeness of fit between the measure and justification? Is it sufficient to demonstrate a reasonable link between a measure and aim, or should the measure be absolutely necessary to achieve said aim? Second, what does it mean for a measure to be necessary (in light of existing alternatives). If strict necessity is mandated, then the measure chosen by the decision-maker must be the only alternative possible. According to Fredman, this question should receive most attention: by requiring a decision-maker to actively look out for alternatives, decision-makers are encouraged to deviate from the status quo, and to better accommodate the needs of those disproportionately impacted. Third, in the assessment of whether the means fit the end, should objective criteria be considered, or does it suffice for the decision-maker to demonstrate that there was a reasonable link between measures and aims? Finally, does the means-end relationship require the establishment of an evidential basis, or can the decision-maker simply assert that a given measure was appropriate? Fredman FBA, *Discrimination Law*, 191–96. In the case of Schaible, the CJEU observed how the principle of proportionality requires that measures do not exceed the limits of what is appropriate and necessary in order to attain the measures pursued. Where there is a choice between measures recourse must be had to the least onerous one; and the disadvantages caused must not be disproportionate to the aims pursued. Case C-101/12 Herbert Schaible v. Land Baden-Wurttemberg ECLI:EU:C:2013:661, § 29 and 76-77.

[731] AG Kokott (31 May 2016) Case C-157/15 Samira Achbita and Centrum voor Gelijkheid van Kansen en voor Racismebestrijding v. G4S Secure Solutions NV., § 112.

[732] Tridimas, *The General Principles of EU Law*, 139.

[733] Tridimas, 173–74.

[734] Tridimas, 74.

In her analysis of the principle of equal treatment's case-law, Croon-Gestefeld observed how the CJEU may abandon its comparative analysis in favour of an approach that is more centred on the standard formula's objective justification prong. In these cases, competing interests are weighed in a thorough and systematised manner, displacing a comparative equality analysis by a careful proportionality review.[735] Rather than "weigh the differences between compared groups with the negative effects on the disadvantaged groups", this proportionality analysis gives the CJEU the possibility to take into account contextual elements that would otherwise be difficult to incorporate as part of a comparative exercise.[736]

A standardised, non-comparative review procedure could prove helpful for the evaluation of algorithmic decision-making systems that differentiate on the basis of seemingly random, opaque, dynamic, or otherwise non-socially significant non-representational differentiation grounds. As aforementioned, given the difficulty of applying the comparability assessment within the algorithmic environment, the CJEU could instead begin its evaluation by examining whether the egalitarian disadvantage a measure is likely to produce for the individuals involved, is outweighed by interests of the decision-maker. In a subsequent step, the Court could further consider whether the differences that define decision-subjects could objectively justify differential treatment. While this approach could initially introduce legal uncertainty over time, the CJEU could establish a more uniform understanding of EU equality, whereby non-discrimination requirements are adapted to take into consideration the changing social and technological environment. To limit the risk of judicial activism, the CJEU's review would need to be clearly structured, as well as consistently and coherently applied.

DIFFERING SCRUTINY AND THE HIERARCHY OF GROUNDS – Within the EU's closed-circuit status and grounds-based system, it if often claimed there is a "hierarchy" or inequality among equality protection: the degree of scrutiny applied by the Court may differ depending on the differentiation ground at play. In other words, the scrutiny applied by the CJEU may fluctuate in accordance with the hierarchy of grounds.[737] Certain grounds enjoy a high degree of protection during the equality review because they are tightly intertwined with EU objectives, such as gender and nationality-

[735] Croon-Gestefeld, *Reconceptualising European Equality Law: A Comparative Institutional Analysis*, 55.

[736] Croon-Gestefeld, 54–55. Tridimas, *The General Principles of EU Law*, 74. Tridimas, 173–74.

[737] For more information, see:Xenidis, 'Transforming EU Equality Law?', 49. Bell, 'The Transformation of EU Anti-Discrimination Law'; Prechal, 'Non-Discrimination Does Not Fall Down From Heaven: The Context and Evolution of Non-Discrimination in EU Law'; Xenidis, 'Transforming EU Equality Law?' See also Bell, 'Article 20: Equality Before the Law', 2021, 609. Croon-Gestefeld, *Reconceptualising European Equality Law: A Comparative Institutional Analysis*, 85.

based discrimination. Other grounds are more protected because they find more coverage within the EU Equality Directives.[738] For instance, in the aforementioned Chez case, the CJEU observed how in cases where there is a difference in treatment on the grounds of racial or ethnic origin, the concept of objective justification must be interpreted strictly.[739] This heightened level of protection could be explained due to the explicit inclusion of race and ethnicity as protected grounds within Article 21 CFR, as well as other international agreements.[740] In addition, racial discrimination would actively undermine EU objectives, such as standard of living and quality of life; maintaining economic and social cohesion and solidarity; and building democratic and tolerant societies that allow the participation of all persons.[741] Other grounds, such as religion, might be met with more restraint by the CJEU due to the cultural and political sensitivities that are involved.[742] Though the application of varying degrees of scrutiny is a legitimate strategy, also within the algorithmic environment, the decision to apply differing levels of scrutiny should be motivated.[743] This is necessary to maintain a clear and well-structured proportionality review.

---

[738] Bell, 'The Transformation of EU Anti-Discrimination Law', 211; Xenidis, 'Transforming EU Equality Law?', 37–55; Belavusau and Henrard, 'A Bird's Eye View on EU Anti-Discrimination Law', 623. Päivi Johanna Neuvonen, '"Inequality in Equality" in the European Union Equality Directives: A Friend or a Foe of More Systematized Relationships between the Protected Grounds?', *International Journal of Discrimination and the Law* 15, no. 4 (December 2015): 222–40, https://doi.org/10.1177/1358229115591412. Xenidis, 'Transforming EU Equality Law?', 37.

[739] CJEU (16 July 2015) Case C-83/14 Chez, ECLI:EU:C:2015:480, § 112, in reference to § 72-74.

[740] CJEU (16 July 2015) Case C-83/14 Chez, ECLI:EU:C:2015:480, § 112, in reference to § 72-73.

[741] CJEU (16 July 2015) Case C-83/14 Chez, ECLI:EU:C:2015:480, § 112, in reference to § 74. Similar statements can be found in the other Equality Directives. For instance, recital 9 of the Framework Equality Directive states how "Employment and occupation are key elements in guaranteeing equal opportunities for all and contribute strongly to the full participation of citizens in economic, cultural and social life and to realising their potential." Whereas, recital 11 states that discrimination based on religion or belief, disability, age or sexual orientation, may undermine the objectives of the EU, including among others raising the standard of living and the quality of life, economic and social cohesion and solidarity."

[742] Xenidis, 'Transforming EU Equality Law?', 49–51. See also: Steve Peers, 'EU Law Analysis: Faith at Work: The CJEU's Headscarf Rulings', *EU Law Analysis* (blog), 17 March 2017, http://eulawanalysis.blogspot.com/2017/03/faith-at-work-cjeus-headscarf-rulings.html.

[743] According to Croon-Gestefeld, the CJEU may sometimes engage in a strict scrutiny approach when interpreting the general principle of equality, among others in cases where discrimination is based upon immutable traits or traits connected to historical disadvantage: "once they have established that they consider one group to be distinct from the other according to one specific criterion, they assess whether this 'otherness', that is to say that criterion of differentiation, justifies the groups differential treatment." The CJEU will move "beyond searching for an objective *reason* for the differentiation and towards scrutinising whether objective *differences* amongst the compared groups justify the differentiation." Because this review contains both an element of comparison and a focus on difference or otherness, it may not sit well with the logic and functioning of the algorithmic environment, where the identified differences between decision-subjects could always be reasonably argued for. Croon-Gestefeld, *Reconceptualising European Equality Law: A Comparative Institutional Analysis*, 51–53.

INSTITUTIONAL SCRUTINY - In cases related to public policy choices, the general principle of equality may be invoked to monitor the consistency and rationality of decision-making.[744] (Public) decision-makers cannot engage in random, inconsistent and arbitrary conduct.[745] The principle of equality simply acts as a democratic guarantee or a "check on the decision-makers": it prevents law makers from imposing differentiation between individuals or groups without good reason, whereby the distinctions made must moreover be reasonably linked to the legitimate purposes pursued by a government.[746] It is said how in these cases, the principle of equal treatment is applied without so-called moral judgment: equality does not endorse a particular theory of equality, nor does it advance a particular idea of what would be the "social good."[747] Instead, equality is compatible with multiple policy goals.[748] Article 20 CFR is often considered as being the Charter representative of the idea of equality as rationality. As an institutional check, Article 20 CFR can find application in highly diverse domains and fields, whereby any differentiation ground, also those that are non-socially salient or non-representational, can become the target of evaluation. The upshot of this "sweeping scope" is the generous margin of appreciation available to public institutions to justify unequal treatment.[749]

According to Bell, when the goal of the CJEU is simply to assess whether a public decision-maker has not acted irrationally, the level of judicial scrutiny applied can be rather light-touch.[750]

Due to its wide scope, Article 20 CFR has been identified as a potential gateway to assess the rightful use of non-protected differentiation grounds in decision-making. Hence, this provision

---

[744] Albeit that certain cases in these areas go beyond a light-touch review. Bell for instance notes how in cases that concern the equal treatment of workers, Article 20 CFR has been applied in a more nuanced manner. See for instance: CJEU (9 March 2017) Case C-406/15 Milkova; ECLI:EU:C:2017:198. See also: Christopher McCrudden, *Buying Social Justice: Equality, Government Procurement, and Legal Change* (Oxford University Press, 2007), https://doi.org/10.1093/acprof:oso/9780199232420.001.0001; McCrudden and Kountouros, 'Human Rights and European Equality Law'.

[745] Croon-Gestefeld, *Reconceptualising European Equality Law: A Comparative Institutional Analysis*, 51. Bell, 'Article 20: Equality Before the Law', 2021, 607; McCrudden, *Buying Social Justice: Equality, Government Procurement, and Legal Change*. Tridimas, *The General Principles of EU Law*, 62.

[746] Tridimas, *The General Principles of EU Law*, 62. In reference to: Jeffrey Jowell, 'Is Equality a Constitutional Principle?', *Current Legal Problems* 47, no. Part_2 (1 January 1994): 7, https://doi.org/10.1093/clp/47.Part_2.1.

[747] However, Tridimas rightfully observes that the evaluation of arbitrary conduct occurs in reference to political ideas and is therefore not politically neutral. Tridimas, *The General Principles of EU Law*, 62; Sottiaux, 'Het Gelijkheidsbeginsel : Langs Oude Paden En Nieuwe Wegen (Artikel, 2008) [WorldCat.Org]', 691.

[748] Tridimas, *The General Principles of EU Law*, 62.

[749] Bell, 'Article 20: Equality Before the Law', 2021, 607.

[750] Bell, 607.

could be valuable for the evaluation of non-representational algorithmic inequalities in particular. One should be careful however when subjugating differential treatment induced by algorithms to a light-level rationality review.

Algorithmically guided institutional decisions might always appear rational due to their data-driven nature. The differentiation they impose can only be meaningfully evaluated, and as a consequence justified, if proper insight is given into how these decisions were built and structured. This need for transparency also applies to the decisional grounds produced on the basis of data-driven analytics.[751] These informational requirements however seem to necessitate a level of scrutiny that goes beyond a mere rationality check: decisions would need to be made explicable and traceable in order to adequately judge whether the output produced by the process was sound, and whether the differences identified between individuals were effectively objective, reasonably justifiable and connected to the purposes pursued. This contextual analysis would be necessary to assess whether the aims and motivations advanced by the decision-maker can be achieved in practice through the specific algorithmic systems that have been put in use.[752] This demonstrates how, during the justification phase sufficient attention must be paid to, and insight given into, the functioning of the algorithmic decision-making structures under review, and not simply end with an assessment of the decision-maker's motivations.[753]

---

[751] See also Section 5.3.3 Rationality, Arbitrariness and Reliability.

[752] According to Croon-Gestefeld, the CJEU's case-law regarding the general principle of equality in EU law, could be roughly categorised into three levels of review, which operate in a continuum and exist side-by side: a low-, high- and intermediate-level review. Whereas the low-level review would correspond to the "rationality" test, under a strict review, the CJEU will evaluate differential treatment in light of the differences between compared groups. As soon as a difference has been established, the CJEU will assess to what extent the differentiation ground justifies differential treatment. Here, the CJEU will move "beyond searching for an objective reason for the differentiation and towards scrutinising whether objective differences amongst the compared groups justify the differentiation." Under a strict scrutiny approach, the CJEU enters into a comparative review in order to weigh the differences between two compared groups against the disadvantages experienced by one of them. Croon-Gestefeld, *Reconceptualising European Equality Law: A Comparative Institutional Analysis*, 50–55.

[753] Sottiaux for example has argued that the introduction of varying levels of intensity during the proportionality analysis could be a valuable structuring mechanism. Such an approach must however clearly delineate what factors will contribute to a higher level of intensity and those that will not. Sottiaux, 'Het Gelijkheidsbeginsel : Langs Oude Paden En Nieuwe Wegen (Artikel, 2008) [WorldCat.Org]', 693.

Open-ended Questions Regarding the Application of the Charter

In accordance with Article 52 (1) CFR, any limitation imposed upon Articles 20 and 21 CFR must in principle be provided for by law and respect their *essence*. Moreover, Article 52 (1) CFR introduced its very own variation of a proportionality test: limitations must be necessary and genuinely meet objectives of general interest or the need to protect the rights and freedoms of others.[754] Questions have therefore been raised to what extent Article 52 (1) CFR's method of analysis, and in particular reference to the concept of the essence, complements or differs from the CJEU's pre-Charter traditional proportionality approach. Muir for example observes how the principle of equality, which has traditionally been understood as an interest weighed against others "does not sit comfortably with an absolutist view of an untouchable essence of fundamental rights."[755] Furthermore, the CJEU's reliance on Article 52 (1) CFR when investigating limitations to the fundamental right to equal treatment appears unclear and inconsistent. In most cases where the general principle of equality applies, the CJEU has continued to make use of the traditional proportionality test.[756] Hence, Muir concludes, "Article 52(1) CFEU is functionally redundant for

[754] See also: CJEU (22 May 2014), Case C-356/12 *Glatzel*, § 42. Article 52 (1) CFR.

[755] Under the absolutist view, the fundamental right to equality would have a core or essence that would not tolerate any restrictions. The notion of essence would be even more problematic if it were to lead to the interpretation that discriminations limited in scope cannot question the essence of equality: "is difficult to reconcile this dignitarian underpinning [of equality]with the idea, put forward in these rulings, according to which restrictions that are limited in scope are for that very reason excluded from the risk of calling into question the foundations of the right to equal treatment." In a similar vein, Wachter has observed that the requirement to have one group be significantly disadvantaged at a far greater number, would exclude protection of algorithmic aggregates. Because algorithm groups are created at a continuous rate, and might not always capture as many people, they might fail to threshold requirements based on numeral criteria. Muir, 'The Essence of the Fundamental Right to Equal Treatment', 822–23 and 838–39. For more information on the notion of "essence": Maja Brkan, 'The Concept of Essence of Fundamental Rights in the EU Legal Order: Peeling the Onion to Its Core', *European Constitutional Law Review* 14, no. 2 (June 2018): 332–68, https://doi.org/10.1017/S1574019618000159; Van Drooghenbroeck and Rizcallah, 'The ECHR and the Essence of Fundamental Rights'; Mark Dawson, Orla Lynskey, and Elise Muir, 'What Is the Added Value of the Concept of the "Essence" of EU Fundamental Rights?', *German Law Journal* 20, no. 6 (September 2019): 763–78, https://doi.org/10.1017/glj.2019.61; Koen Lenaerts, 'Limits on Limitations: The Essence of Fundamental Rights in the EU', *German Law Journal* 20, no. 6 (September 2019): 779–93, https://doi.org/10.1017/glj.2019.62. Wachter, 'Affinity Profiling and Discrimination by Association in Online Behavioral Advertising', 44–46.

[756] Muir, 'The Essence of the Fundamental Right to Equal Treatment'. Muir refers to the differing approach that is present in (among others) the following the cases of Glatzel, Leger and Fries: CJEU (22 May 2014), Case C-356/12 *Glatzel*; CJEU (29 April 2015) C-528/13 Léger, ECLI:EU:C:2015:288; CJEU (5 July 2017) Case C-190/16 Werner Fries v. Lufthansa CityLine GmbH, ECLI:EU:C:2017:513.

the analysis of restrictions to the fundamental right to equal treatment, and the concept of 'essence' can simply be subsumed in the traditional proportionality test."[757]

An additional issue that has not yet been resolved is whether the Charter maintains the distinction between direct and indirect discrimination. In certain cases, the CJEU has established there was discrimination without specifying however whether a challenged piece of legislation was directly or indirectly discriminatory.[758] In traditional EU Law however, the identification of a measure as directly or indirectly discriminatory is often a crucial component of the justification phase. As aforementioned, under the EU Equality Directives, direct forms of discrimination are typically prohibited whereas indirectly discriminatory measures can be objectively justified. If the application of the Charter would necessitate the performance of a proportionality review, as Article 52(1) CFR seemingly mandates, direct forms of discrimination, even when based upon suspect grounds, could become justifiable under certain conditions, for instance in cases where the EU Equality Directives do not apply. Would this however be a desired approach in relation to grounds the use of which is often considered morally reprehensible?[759]

5.3.3 Normative Guidance?

In its interpretative function, the CJEU, with the support of its Advocate Generals, provides additional (normative) substance to the principle of equality via its analysis of the Charter and the EU Equality Directives. Unfortunately, the EU's case-law is not necessarily hallmarked by explicit normative analyses. Within the EU's grounds-based system, normative guidance would nonetheless be welcome. It would enable decision-makers to foresee under what conditions the use of a given (protected) differentiation ground could be justified.

---

[757] Muir, 818.

[758] See for instance CJEU (11 April 2013) C-401/11 Blanka Soukupová v. Ministerstvo zemědělství; ECLI:EU:C:2013:223.

[759] As Kilpatrick and Eklund observe, "this criss-crossing between methods of justification under the auspices of the Charter [would be] to the detriment of those people it sets out to protect." Eklund and Kilpatrick, 'Article 21 EU Charter of Fundamental Rights', 634–35. See also: CJEU (29 April 2015) Case C-528/13 Léger; Ward, 'The Impact of the EU Charter of Fundamental Rights on Anti-Discrimination Law'. Evelyn Ellis and Philippa Watson, 'Key Concepts in EU Anti-Discrimination Law', in *EU Anti-Discrimination Law*, 2nd ed. (Oxford: Oxford University Press, 2012), https://doi.org/10.1093/acprof:oso/9780199698462.003.0004.

AG Sharpston observed how "the criteria of relevant resemblances and differences vary with the *fundamental moral outlook* of a given person or society [own emphasis added]."[760] Under this view, equality emerges as an evolutive concept, whereby the difference between acceptable differential treatment and unacceptable discrimination should not be found in some universal principle or theory, but derived from people's social acceptance of a certain characteristic as relevant or arbitrary for decision-making.[761] The guiding value of such an "empirical" approach might lack specificity though: how can we uncover the position people hold vis-à-vis the use of certain differentiating criteria? Others have noted how, in light of the requirement to ensure the uniform application of EU law, a common understanding of equality (across grounds) would be desirable.[762] Xenidis for example has similarly argued that once the conceptual and normative foundations of equality are articulated and linked to each ground of protection, non-discrimination rights will be able to reach their full protective potential and transformative impact.[763]

In this Section, I therefore analyse the case-law in order to delve deeper into the normative underpinnings underlying the EU's concept of equality. These rationales can help clarify the heightened level of scrutiny applied by the CJEU, or further flesh out the egaliatarian objectives pursued by the EU at large. As was the case within the ECtHR's approach towards equality, these rationales are often recognised and defined in relation to the EU's protected discrimination grounds. Still, I argue that it is possible to view them more independently: to identify additional (normative) anchoring points that ground the notion of equality when it is positioned against competing interests. This may enable us to unearth other defining components that shape the EU conceptualisation of equality, which may be important when the notion is applied to evaluate algorithmic inequalities in non-distributive contexts. In addition, if the CJEU were to ever expand

---

[760] Opinion of AG Sharpston (22 May 2008) Case-427/06 Birgit Bartsch v Bosch und Siemens Hausgeräte (BSH) Altersfürsorge GmbH; ECLI:EU:C:2008:297; § 49, § 44. In reference to H.L.A. Hart's The Concept of Law (2nd Edition, 1994), p. 159-163. See also: Bell, 'Article 20: Equality Before the Law', 2021, 605.

[761] Opinion of AG Sharpston (22 May 2008) Case-427/06 Birgit Bartsch v Bosch und Siemens Hausgeräte (BSH) Altersfürsorge GmbH; ECLI:EU:C:2008:297, § 46-56.

[762] Lucy Vickers, 'ECJ Headscarf Series (2): The Role of Choice; and the Margin of Appreciation', *Strasbourg Observers* (blog), 8 September 2016, https://strasbourg.weichie.dev/2016/09/08/blog-series-the-role-of-choice-and-the-margin-of-appreciation/. In reference to: CJEU (11 July 2006) Case C-13/06 Sonia Chacón Navas v. Eurest Colectividades SA *(Chacón Navas)* ; and juxtaposing AG Sharpston and Kokott's opinions in the Bouganoui and Achbita cases. Opinion of AG Sharpston (13 July 2016) Case C-188/15 Asma Bougnaoui and ADDH v Micropole SA; Opinion of AG Kokott (31 May 2016) Case C-157/15 Samira Achbita and Centrum voor Gelijkheid van Kansen en voor Racismebestrijding v. G4S Secure Solutions NV.

[763] Xenidis, 'Transforming EU Equality Law?', 56.

the Charter's protected scope to non-protected grounds, these anchoring points could be evaluated in light of their ability to capture the harm associated with their use. In this context, they can be used as a forward-looking purposes: would they allow the evolution of equality into a principle that can withstand social and technological change? We can ask ourselves the question: do these rationales make sense? Not only in their ability to capture existing social injustice, but also in their capacity to tackle new harms. In addition, do they make sense in relation to protected grounds only, or do they refer to harms that should be of broader concern?

Like the ECtHR's conceptualisation of equality, it is possible to discern a distinction between rationales that are rooted within a formal approach towards equality and rationales that are more substantive in nature. Whereas the former tend to focus on the nature of the differentiation ground at play (e.g., the irrelevance or immutability of the differentiating trait), the latter are more sensitive to the social context that allows such disadvantage to arise (e.g., existing prejudice and stereotypes towards certain vulnerable (social) groups). Under both rationales, the particular egalitarian disadvantage experienced by those who face discrimination has been further specified as also comprising a violation of (individual) dignity and autonomy. Sometimes the evaluation of discriminatory acts is assessed through the lens of multiple rationales. As a consequence, depending on the differentiation ground and context at play, differential treatment may be more easily, or conversely, more unlikely, to find justification.[764] Given that these rationales are similar to those found under ECtHR case-law, for their application onto the algorithmic environment, reference can be made to the corresponding Sections in the previous Chapter.[765] Still, this Section will conclude with a more in-depth investigation of notions (primarily) associated with the use of equality as an institutional or infrastructural check within the EU.[766] Here, non-suspect differentiation grounds can be made subject to an evaluation and assessed in terms of their rationality, arbitrariness and, in the case of algorithms, reliability.

---

[764] At the same time, the EU also maintains a "hierarchy" or inequality among equality protection. Certain grounds enjoy a high degree of protection because they tightly intertwined with EU objectives, such as gender and nationality-based discrimination. Other grounds have more protection because they find more coverage within the EU Equality Directives. Bell, 'The Transformation of EU Anti-Discrimination Law', 211; Xenidis, 'Transforming EU Equality Law?', 37–55; Belavusau and Henrard, 'A Bird's Eye View on EU Anti-Discrimination Law', 623. Neuvonen, '"Inequality in Equality" in the European Union Equality Directives'. Xenidis, 'Transforming EU Equality Law?', 37.

[765] Chapter 4, Section 4.3.2 Normative Guidance.

[766] It should be noted however that the protected grounds of Article 21 (1) CFR may also function as an institutional check.

Dignity and Autonomy

Dignity and autonomy are the anchoring points that ground and give normative substance to EU equality laws.[767] In his opinion to the case of *Coleman*, AG Maduro clarified the foundational importance dignity and autonomy hold as values underlying a society's commitment to equality, and the protection of people against discrimination.

First, human dignity, at its bare minimum, entails the recognition of every individual's equal moral worth. Human dignity requires people and political actors to value the "intrinsic importance of every human life", and they should not act in ways that denies the intrinsic importance of every human life.[768] Second, people are autonomous when they are "able to design and conduct the course of their lives through a succession of choices among different valuable options."[769] AG Maduro thus understands autonomy as the freedom a person has over a range of options to choose and develop their life's goals, give structure to their personal integrity, their sense of dignity and self-respect: "a commitment to autonomy means that people must not be deprived of valuable options in areas of fundamental importance for their lives by reference to suspect classifications."[770] Inequality is unjustified where a person's freedom of choice is curtailed and shaped by prejudice. People have an equal right towards their self-authorship which can be maintained by guaranteeing people's equality within fundamental areas of importance. This account reflects the EU ideal whereby discrimination is wrong not only because it can have a negative impact on the socio-economic outcomes of people, but also on a more foundational level, on the moral and socio-relational position they hold.[771]

---

[767] Prechal, 'Non-Discrimination Does Not Fall Down From Heaven: The Context and Evolution of Non-Discrimination in EU Law'.

[768] AG Maduro Opinion (31 January 2008) Case C-303/06 of S. Coleman v. Attridge Law and Steve Law (*Coleman)*, § 9.

[769] AG Maduro Opinion (31 January 2008) Case C-303/06 of S. Coleman v. Attridge Law and Steve Law (*Coleman)*, § 9. In reference to Joseph Raz, The Morality of Freedom. This autonomy account has also been used to ground the fundamental right to privacy. See for example:

[770] AG Maduro Opinion (31 January 2008) Case C-303/06 of S. Coleman v. Attridge Law and Steve Law (*Coleman)*, § 9-11. In reference to Joseph Raz, The Morality of Freedom.

[771] See also AG Sharpston's obvservations concerning AG Maduro's account of Coleman in the case of Bougnaoui. AG Shaprston agrees with Maduro and notes that this account emphasises "that discrimination has both a financial impact (because it may touch on a person's ability to earn a living in the employment market) and a moral impact (because it may affect that person's autonomy). I would add that anti-discrimination legislation must, in the same way as all other legislation, be applied in a way that is effective. It must also be applied in accordance with established principles." Opinion of AG Sharpston (13 July 2016) Case C-188/15 Asma Bougnaoui, § 71 and 72.

According to AG Maduro, the goal of EU equality laws is to protect the dignity and autonomy of persons belonging to a suspect group. The most obvious ways in which a person's dignity and autonomy can be affected, is to target them directly due to their possession of a suspect characteristic.[772] The reason why these grounds enjoy a heightened level of protection, and why their use is considered a particular threat towards people's dignity and autonomy, can be further linked to reasons that are either formal or substantive in nature (or both).[773]

Irrelevance and Immutability

Under a formal approach towards equality, suspect classifications enjoy a heightened degree of protection due to their irrelevance or immutability. Either a differentiation ground is not considered relevant for the decision at hand, or the differentiation ground cannot be used because people are unable to exercise (a required degree of) control over these traits.

IRRELEVANCE - In the same opinion, AG Maduro argued that to treat someone lesser because of their religious beliefs, age, disability, and sexual orientation, etc. is to directly undermine that person's moral worth. Conversely, if we want to recognise people's equal moral worth "we should be *blind* to considerations of this type when we impose a burden on someone or deprive someone of a benefit (own emphasis added)." This statement reflects the irrelevance rationale. These traits should be disregarded, or "put differently, [they] should not play any role in any assessment as to whether it is right or not to treat someone less favourably."[774]

IMMUTABILITY - In the *Achbita* case, AG Kokott reviewed the proportionality (*sensu stricto*) of a workplace prohibition to wear religious symbols. Here, the AG argued that "unlike sex, skin colour, ethnic origin, sexual orientation, age or a person's disability, the practice of religion is not so much

---

[772] If dignity and personal autonomy are the underlying value that determine the level of protection afforded by equality and non-discrimination law, it should extend to more situations than only those where a "suspect classification" lies at the basis of differential treatment. People can also be deprived of valuable life opportunities when they are judged on the basis of non-protected classifications. This was illustrated among others by the deployment of educational algorithms, where students were deprived of their opportunity to enjoy higher education on the basis of historical school results. The argument would need to be made that dignity and autonomy could be linked or limited to suspect criteria only. However, it seems more reasonable to assume that dignity and autonomy should not be limited in this way.

[773] AG Maduro Opinion (31 January 2008) Case C-303/06 of S. Coleman v. Attridge Law and Steve Law (*Coleman)*, § 10, read in combination with § 14.

[774] AG Maduro Opinion (31 January 2008) Case C-303/06 of S. Coleman v. Attridge Law and Steve Law (*Coleman)*, § 10, readin in combination with 14.

an unalterable fact as an aspect of an individual's private life, and one, moreover, over which the employees concerned can choose to exert an influence."[775] One can moderate the exercise of religion, whereas one cannot moderate their sex, skin colour, ethnicity, sexual orientation, age or disability.[776] Under the immutability rationale, the degree of control or choice people have over certain characteristics can be taken into account during the proportionality assessment.

As aforementioned however, both rationales can undermine the substantive goals EU equality laws look to achieve to protect marginalised or otherwise vulnerable groups. The irrelevance rational aims to make abstraction of certain identity-markers within decision-making.[777] Yet, as Fredman observed, an "individual's social, economic, and political situation is still heavily determined by these very characteristics."[778] The irrelevance rationale therefore fails to understand the particular importance certain characteristics, and associated group-affinities, hold for people. By focusing on the nature of the traits, one foregoes the social implications of being treated because of the possession of a given trait. First, a "blind" irrelevance approach risks to decontextualise the disadvantage certain groups face due to their possession of a given characteristic. Second, it underestimates how certain characteristics can have an impact on one's identity. In this context, the differences between people should not be rendered invisible, but instead accommodated by eliminating the (socio-relational and distributional) disadvantage attached to the difference.[779] Where immutability is relied upon, people are seen as rational and autonomous actors that are in full control over their life's options. The choices people make, which could also include people's choice to wear certain religious symbols, are often made under social conditions over which they hold no, little or less control.[780] These conditions might be overlooked where people are considered as having full agency and control over certain characteristics, and as

[775] Opinion of AG Kokott (31 May 2016) Case C-157/15 Samira Achbita and Centrum voor Gelijkheid van Kansen en voor Racismebestrijding v. G4S Secure Solutions NV. , § 116. See also: Vickers, 'ECJ Headscarf Series (2)'. This can be contrasted with AG Sharpston's opinion in the headscarf case of Bougnaoui. Here, the dignity and autonomy component were again emphasized. was placed Opinion of AG Sharpston (13 July 2016) Case C-188/15 Asma Bougnaoui, § 71 and 72.

[776] Opinion of AG Kokott (31 May 2016) Case C-157/15 Samira Achbita and Centrum voor Gelijkheid van Kansen en voor Racismebestrijding v. G4S Secure Solutions NV. , § 116.

[777] See above Chapter 3, Section 3.1.2 "Formal and Substantive Approaches towards Equality".

[778] Fredman, 'Substantive Equality Revisited', 719.

[779] Fredman, 720.

[780] See also observations above in Chapter 4, Section 4.3.2 "Normative Guidance" and "Immutability". Timmer and Peroni, 'Bah v UK: On Immigration, Discrimination and Worrisome Reasoning'. Solanke, *Discrimination as Stigma*, chaps 2, Legal Protection from Discrimination.

a consequence, once a trait has been labelled as "immutable", we might overly punish people for their possession of them.[781]

Under both rationales, it is believed that certain traits have a particular (and often absolute) nature which renders their use prohibited. This narrow focus might disregard the intricate relationship the possession of a given characteristic has with the particular (socio-contextual) disadvantage certain groups experience. While one should indeed remain careful to treat people differently on the basis of the protected characteristics, by prohibiting their use solely because of their (supposed) nature, one might eliminate the possibility of harnessing these differences positively as a means to achieve social change, i.e., to understand and accommodate the disadvantage associated with them.

Anti-Stereotyping and Vulnerability

Irrelevance and immutability have been highlighted as rationales that clarify the heightened protection of suspect characteristics. These traits have also gained a symbolic value: due to their possession of a given trait, certain social groups face prejudice, stereotypes, or other vulnerabilities.

ANTI-STEREOTYPING – EU equality laws have tried to counteract the adverse social effects that stem from stereotypical attitudes.[782] The CJEU has recognised how inequality may result from 'social attitudes, behaviour and structures.'[783] For example, the Court has confirmed that a differentiating measure should not have an offensive or stigmatising impact.[784] The CJEU has further argued that EU law precludes national provisions that are "liable to perpetuate a traditional distribution of the roles of men and women by keeping men in a role subsidiary to that of women in relation to the exercise of their parental duties."[785] The CJEU also confirmed that differentiating measures should

[781] Fredman, 'Substantive Equality Revisited', 738.

[782] AG Alber Opinion (19 March 2002) Case C-476/99 H. Lommers v. Minister van Landbouw, Natuurbeheer en Visserij, § 36. In reference to CJEU (11 November 1997) Case C-409/95 of Helmut Marschall v. Land Nordrhein-Westfalen, § 29 and 31; CJEU See also: CJEU (19 March 2002) Case C-476/99 H. Lommers v. Minister van Landbouw, Natuurbeheer en Visserij.

[783] CJEU (17 October 1995) Case C-450/93 of Eckhard Kalanke v. Freie Hansestadt Bremen, § 20; CJEU (11 November 1997) Case C-409/95 of Helmut Marschall v. Land Nordrhein-Westfalen, § 28.

[784] CJEU (16 July 2015) Case C-83/14 Chez

[785] CJEU (16 April 2015) Case C-222/14 Konstantinos Maïstrellis v. Ypourgos Dikaiosynis, Diafaneias kai Anthropinon Dikaiomaton, § 50; CJEU (30 September 2010) Case C-104/09 Pedro Manuel Roca Álvarez v. Sesa Start España ETT SA, § 36. Timmer, 'Gender Stereotyping in the Case Law of the EU Court of Justice', 42.

also avoid the reinforcement of traditional structures even if the purported aim the measures involved is to abolish *de facto* inequality.[786] In the case of *Achbita,* AG Kokott observed how Article 21 CFR and the EU Equality Directives seek to eliminate traditional prejudice and break with outdated structures: "Under no circumstances, therefore, would it be permissible to perpetuate existing differences of treatment simply because doing so accords with certain traditions, customs or social structures."[787] In order to break down inequality linked to existing social structures, a proportionality review could take into account the representational and socio-relational disadvantage a differentiating measure can impose.

VULNERABILITY - For certain (social) groups, the possession of a protected characteristic, such as a person's disability, signifies social "vulnerability." This too can explain why their use as a differentiation ground should be approached with more caution. In addition, the vulnerability of social groups clarifies the problematic nature of discrimination through association. In the case of *Coleman*, AG Maduro observed how certain social groups are more vulnerable and have come to rely on others with whom they are associated in order to lead a dignified and autonomous life in accordance with their fundamental choices.[788] Discrimination by association is troublesome in a twofold way. On the one hand, these discriminatory acts interfere with the dignity and autonomy of those directly targeted by the differentiating measure. On the other hand, their discrimination further undermines the dignity and autonomy of those belonging to a protected or vulnerable

---

[786] CJEU (19 March 2002) Case C-476/99 H. Lommers v. Minister van Landbouw, Natuurbeheer en Visserij, § 41. This difficulty however might also be faced by the CJEU. Timmer has observed that the CJEU sometimes reinforces and at other times contests the use of stereotypes. For instance, the Court has been criticised for perpetuating gender stereotypes in the areas of pregnancy and maternity protection. For more information on the CJEU's anti-stereotyping approach, see primarily: Timmer, 'Gender Stereotyping in the Case Law of the EU Court of Justice'. On traditional gender views concerning maternity leave, see: Case C 184/83 (12 July 1984) Ulrich Hofmann v Barmer Ersatzkasse, ECLI:EU:C:1984:273. See also: Miguel De La Corte-Rodríguez, 'The First Revision of the Hofmann Case Law on Maternity Leave and Discrimination Against Fathers: Care-Giving as the Pivot?', *Industrial Law Journal* 50, no. 2 (1 June 2021): 306–23, https://doi.org/10.1093/indlaw/dwab008.

[787] AG Kokott Opinion (31 May 2016) Case C-157/15 Samira Achbita and Centrum voor gelijkheid van kansen en voor racismebestrijding v. G4S Secure Solutions NV, § 123. Still, both AG Kokott's oinion, as well as the case-law of the Court with regard to religious stereotypes and prejudice, and in particular with regard to the wearing of political and religious symbols, have been criticized as they would "legitimize and even provide a recipe for discrimination of employees on the basis of their religious or other conviction." Saïla Ouald-Chaib and Valeska David, 'European Court of Justice Keeps the Door to Religious Discrimination in the Private Workplace Opened. The European Court of Human Rights Could Close It.', *Strasbourg Observers* (blog), 2017, https://strasbourg.weichie.dev/2017/03/27/european-court-of-justice-keeps-the-door-to-religious-discrimination-in-the-pri
vate-workplace-opened-the-european-court-of-human-rights-could-close-it/.

[788] AG Maduro Opinion (31 January 2008) Case C-303/06 of S. Coleman v. Attridge Law and Steve Law (*Coleman)*, § 14.

group. As AG Maduro explains: "When the discriminator deprives an individual of valuable options in areas which are of fundamental importance to our lives because that individual is associated with a person having a suspect characteristic then it also deprives that person of valuable options and prevents him from exercising his autonomy. Put differently, the person who belongs to the suspect classification is excluded from a range of possibilities that would otherwise have been open to him."[789]

These substantives rationales articulate discrimination's problematic nature in relation to the (socio-relational) disadvantage they impose onto decision-subjects. Where the detrimental character of differentiation is located in the nature of the trait, equality is framed primarily in negative terms in the sense that once a trait has been identified as either irrelevant or immutable it should not be used. Under a substantive rationale, equality's positive function becomes more clearly visible. It is a tool for breaking down social structures and accommodating the needs of those most vulnerable. As illustrated by the previous paragraphs, in the cases of *Coleman* and *Achbita,* the AGs relied on both types of rationales to clarify the functioning of EU Equality Law. Under this hybrid approach however, the formal approach might constrain the substantive reach or goals pursued by EU equality law as the former cannot fully capture the particular disadvantage the latter tries to address.

Rationality, Arbitrariness and Reliability

In this final Section, I will address three (interrelated) notions that inform the equality debate. The first two notions are *rationality* and *arbitrariness*. They are typically associated with equality's function as an institutional or infrastructural benchmark, and are not normative *per se*.[790] Instead, they can be used to describe the soundness of the differentiation ground. In this context, the soundness of a differentiation ground might also be informed by a third notion, its *reliability*.

RATIONALITY AND ARBITRARINESS – As above-mentioned, in certain cases, equality acts as a catch-all principle that is applicable in a wide variety of circumstances and a wide variety of grounds.[791] Adherence to the general principle of equality requires that a decision-making process is rational. Differential treatment should be justified in accordance to "consistently applied, persuasive and

[789] AG Maduro Opinion (31 January 2008) Case C-303/06 of S. Coleman v. Attridge Law and Steve Law (*Coleman)*, § 14.

[790] This model of equality tends to be grounded in a formal understanding of equality too.

[791] Albeit that as of now, this reading of equality is only applicable in the "public realm, broadly defined." McCrudden and Kountouros, 'Human Rights and European Equality Law', 75.

acceptable criteria."[792] This criterion typically prohibits a decision-maker from engaging into random, arbitrary or inconsistent behaviour and can be highly deferential to decision-makers.[793] There is no concrete definition of "arbitrariness" however.[794] According to AG Sharpston, arbitrary differential treatment stands opposed to justifiable differential treatment: "The difference between (acceptable) differential treatment and (unacceptable) discrimination lies not in whether people are treated differently, but in whether society accepts as justifiable the criteria whose application results in different treatment, or whether, on the contrary, they are considered as arbitrary." Yet, this reading remains too broad: we might find that the use of a decision ground cannot be justified, without necessarily believing that ground was random. The German Federal Constitutional Court developed an "arbitrariness formula.[795] Under this test, a measure will be in violation of equality when there is no reasonable ground, arising from the nature of things, or any other intelligible or objectively plausible reason for differentiation.[796] Differential treatment would then be arbitrary if entirely similar cases are treated differently in a manner that is clearly irrelevant or unobjective.[797]

Creel and Hellmann see three different understandings of arbitrary decisions: a) unpredictable decisions, b) decisions unconstrained by ex-ante rules or standards and c) decisions that are nonsensical or unreasonable.[798] Having their basis in data, the decisional knowledge derived from

[792] McCrudden and Prechal, 'The Concepts of Equality and Non-Discrimination in Europe: A Practical Approach', 12.

[793] In certain jurisdictions however, the justification procedure is stricter. See for instance: McCrudden and Prechal, 14–15; Croon-Gestefeld, *Reconceptualising European Equality Law: A Comparative Institutional Analysis*, 55–58. See also: Bell, 'Article 20: Equality Before the Law', 2021, 607.

[794] Opinion of AG Sharpston (22 May 2008) Case-427/06 Birgit Bartsch v Bosch und Siemens Hausgeräte (BSH) Altersfürsorge GmbH; ECLI:EU:C:2008:297, § 56.

[795] The arbitrariness formula reads: "the equality guarantee is violated when a reasonable ground, that is due to the nature of things or otherwise objectively plausible, cannot be found for the legal differentiation or equal treatment; in short, if the provision must be called arbitrary. (Translation provided by Croon-Gestefeld). This arbitrariness formula was however replaced by a proportionality analysis, where a heightened level of scrutiny is applied (among others) where different groups of persons are treated unequally.

[796] McCrudden and Prechal, 'The Concepts of Equality and Non-Discrimination in Europe: A Practical Approach', 14. Croon-Gestefeld, *Reconceptualising European Equality Law: A Comparative Institutional Analysis*, 56.

[797] McCrudden and Prechal, 'The Concepts of Equality and Non-Discrimination in Europe: A Practical Approach', 14. Croon-Gestefeld, *Reconceptualising European Equality Law: A Comparative Institutional Analysis*, 56.

[798] Kathleen Creel and Deborah Hellman, 'The Algorithmic Leviathan: Arbitrariness, Fairness, and Opportunity in Algorithmic Decision-Making Systems', *Canadian Journal of Philosophy* 52, no. 1 (January 2022): 28, https://doi.org/10.1017/can.2022.3.

and applied through analytical processes may always appear rational and justified.[799] Algorithms have the capacity to render every differentiation ground non-arbitrary and relevant; from a decision-maker's perspective this is often their entire purpose.[800] The default rationality ascribed to data-driven systems must therefore be actively argued against in order for an evaluation of their functioning to be effective.[801] At the same time, Creel and Hellman believe that the arbitrariness of algorithms is not necessarily the problem, but the systematic exclusions these systems produce are. In this case the rationality of a given measure should not be assessed with a light touch, but through a balancing act, whereby every decision-making act is critically evaluated not only in light of the decisional criteria that were used, but also the harms produced.

RELIABILITY – In certain cases, the CJEU had to rule on the discriminatory nature of certain data for profiling purposes.[802] In these cases, the CJEU shed more light onto the relationship it sees between the reliability of information used and discrimination.[803] Whereas a full analysis and critique of this case-law would is outside of the scope of this dissertation, some elements will be highlighted as they contextualise among others the CJEU's approach towards equality in an automated setting, and the supportive function data protection safeguards have in this regard.[804]

---

[799] For a more in-depth discussion of the notion of arbitrariness in relation to algorithmically guided decision-making: Creel and Hellman, 27.

[800] Wachter, 'The Theory of Artificial Immutability', 20–21.

[801] It could be argued however that judges, when doing so, would actively enter into the political landscape. They would assess when, or formulate conditions under which, algorithmic differentiation grounds would be non-arbitrary. While this appears as a form of judicial activism, it must be observed that in other contexts the arbitrariness of a given conduct is often evaluated in reference to political ideas, such as egalitarianism. As such, one could also take into account the egalitarian harms produced by a system. Tridimas, *The General Principles of EU Law*, 62.

[802] See in relation to the processing of PNR Data, and the interpretation of Directive (EU) 2016/68: CJEU (21 June 2022), Case C-817/19 Ligue des Droits Humains v. Conseils des Ministres; ECLI:EU:C:2022:491; CJEU (26 July 2017) Opinion 1/15 of the Court on the transfer of Passenger Name Record Data from the European Union to Canada, ECLI:EU:C:2017:592. In relation to profiling on the basis of electronic communications data; including traffic and location data, and the interpretation of Directive 2002/58/EC: See also: CJEU (5 April 2022) Case C-140/20 G.D. v. Commissioner of An Garda Siochana; ECLI:EU:C:2022:258; CJEU (6 October 2020) La Quadrature du Net and Others v Premier Ministre and Others; ECLI:EU:C:2020:791; and CJEU (21 December 2016) Joined Cases C-203/15 and C-698/15, Tele2 Sverige AB, ECLI:EU:C:2016:970.

[803] Of course, the notion of discrimination is here understood mainly in reference to protected characteristics.

[804] For a critique, see among others: Maria Tzanou and Spyridoula Karyda, 'Privacy International and Quadrature Du Net: One Step Forward Two Steps Back in the Data Retention Saga?', SSRN Scholarly Paper (Rochester, NY, 11 July 2021), https://doi.org/10.2139/ssrn.3970756.

In its Opinion 1/15 on the Draft Agreement between Canada and the EU for the transfer of Passenger Name Record Data, the CJEU had to evaluate the use of risk assessment methods of air passengers through automated processing techniques.[805] These risk assessments are carried out on the basis of unverified personal data and based upon pre-established models and criteria, and necessarily present a margin of error.[806] Those models and criteria, the CJEU observed, should be specific, reliable and non-discriminatory. Whereas models and criteria must be specific and reliable in order to make certain that only individuals presenting a "reasonable suspicion" towards a threat are targeted, these requirements also contribute to the non-discriminatory nature of automated decision-making processes: "in order to ensure that […] the pre-established models and criteria, the use that is made of them and the databases used are not discriminatory and are limited to that which is strictly necessary, the reliability and topicality of those pre-established models and criteria and databases used should, taking account of statistical data and results of international research, [be subject to review]."[807] This linkage between reliability and discrimination is also present within the GDPR, where data controllers are requested to use appropriate mathematical and statistical procedures for profiling and implement appropriate

[805] PNR data can be incredibly sensitive, as is noted by the CJEU: "Even if some of the PNR data, taken in isolation, does not appear to be liable to reveal important information about the private life of the persons concerned, the fact remains that, taken as a whole, the data may, inter alia, reveal a complete travel itinerary, travel habits, relationships existing between air passengers and the financial situation of air passengers, their dietary habits or state of health, and may even provide sensitive information about those passengers." Opinion 1/15 of the Court on the transfer of Passenger Name Record Data from the European Union to Canada, ECLI:EU:C:2017:592, § 128.

[806] Opinion 1/15 of the Court on the transfer of Passenger Name Record Data from the European Union to Canada, ECLI:EU:C:2017:592, § 169. It should be noted however that in this particular case, no reference was made to Article 20 or 21 CFR. See also: La Quadrature du Net and Others v Premier Ministre and Others; ECLI:EU:C:2020:791, § 189. In the latter case, similar conditions were imposed onto the retention of traffic and location data: "since the automated analyses of traffic and location data necessarily involve some margin of error, any positive result obtained following automated processing must be subject to an individual re-examination by non-automated means before an individual measure adversely affecting the persons concerned is adopted, such as the subsequent real-time collection of traffic and location data, since such a measure cannot be based solely and decisively on the result of automated processing. Similarly, in order to ensure that, in practice, the pre-established models and criteria, the use that is made of them and the databases used are not discriminatory and are limited to that which is strictly necessary in the light of the objective of preventing terrorist activities that constitute a serious threat to national security, a regular re-examination should be undertaken to ensure that those pre-established models and criteria and the databases used are reliable and up to date."

[807] Opinion 1/15 of the Court on the transfer of Passenger Name Record Data from the European Union to Canada, ECLI:EU:C:2017:592, § 172.

technical and organisational measures in order to lower the presence of inaccuracies, errors and discriminatory effects on natural persons.[808]

In order to reduce the margin of error, non-automated, individual re-examination mechanisms should be adopted.[809] In this context, protection standards that stem from privacy and data protection obligations, such as the need to have a human-in-the-loop, might also benefit the protection against discrimination. Further specifications were provided by the CJEU in their interpretation of the PNR Directive, which allows among others the comparing of PNR data against databases and the automated processing of PNR data on the basis of pre-determined criteria, similarly as the EU-Canada Agreement.[810] According to the PNR Directive, the assessment of passengers against pre-determined criteria must be carried out in a manner that is non-discriminatory.[811] Those criteria must be targeted, proportionate, specific and subject to a periodic review. In no circumstances can they be based upon a person's race or ethnic origin, political opinions, religion or philosophical beliefs, trade union membership, health, sexual life, or sexual orientation, or can they place persons having those protected characteristics at a particular disadvantage.[812]

---

[808] Recital 71 GDPR. More specifically, "discriminatory effects on the basis of racial or ethnic origin, political opinion, religion or beliefs, trade union membership, genetic or health status or sexual orientation, or that result in measures having such an effect." It can also be noted that Article 9 GDPR installs a principled prohibition on the processing of special categories of data, i.e. "personal data revealing racial or ethnic origin, political opinions, religious or philosophical beliefs, or trade union membership, and the processing of genetic data, biometric data for the purpose of uniquely identifying a natural person, data concerning health or data concerning a natural person's sex life or sexual orientation." The latter are typically associated with equality and non-discrimination.

[809] Opinion 1/15 of the Court on the transfer of Passenger Name Record Data from the European Union to Canada, ECLI:EU:C:2017:592, § 173. See also : CJEU (21 June 2022), Case C-817/19 Ligue des Droits Humains v. Conseils des Ministres; ECLI:EU:C:2022:491 : § 178. The need for a human in the loop can be found in both the GDPR and the PNR Directive: Article 22 (3) GDPR as well as Article 6 (5) PNR Directive. Directive (EU) 2016/681 of the European Parliament and of the Council of 27 April 2016 on the use of passenger name record (PNR) data for the prevention, detection, investigation and prosecution of terrorist offences and serious crime. L 119/132

[810] See also CJEU (21 June 2022), Case C-819/19 Ligue des Droits Humains v. Conseils des Ministres; ECLI:EU:C:2022:491.

[811] Article 6 (4) PNR Directive.

[812] However, according to the CJEU these requirements not only apply to the processing of PNR data against pre-determined criteria (Article 6 (b) PNR Directive), but also where PNR are compared against databases relevant for the purposes of fighting serious crime (Article 6 (a) PNR Directive): "If, by referring to Article 6(3)(b) of that directive, the wording of Article 6(4) thereof covers only the processing of PNR data against pre-determined criteria, the latter provision must be interpreted, in the light of Articles 7, 8 and 21 of the Charter, as meaning that the requirements it lays down apply *mutatis mutandis* to the comparison of those data against the databases referred [to in Article 6(3) a PNR Directive], especially since those requirements correspond, in essence, to those adopted by the case-law arising from

With regard to these criteria, the CJEU first established that the need for them to be pre-determined precludes "the use of artificial intelligence technology in self-learning systems, capable of modifying without human intervention or review the assessment process and, in particular, the assessment criteria on which the result of the application of that process is based as well as the weighting of those criteria."[813] Self-learning systems, argues the CJEU, might render impossible the capacity to individually review and monitor the lawfulness of a decision, *among others* due to complexity and opacity they exhibit. On this point, AG Pitruzzella specified that algorithms should operate in a manner that is transparent in the sense that (the result of the) decision-making process should be traceable and identifiable.[814] The latter, the CJEU adds, would deprive data subjects of their right to challenge the non-discriminatory nature of the results obtained.[815] Clear and precise guidelines and support, including objective review criteria, during the individual, non-automated review could moreover help ensure respect for Article 21 (1) CFR, and guarantee uniformity in the assurance thereof.[816] Likewise, self-monitoring and documentation efforts can help contribute to the lawfulness, and in particular, non-discriminatory

---

Opinion 1/15 *(EU-Canada PNR Agreement)* of 26 July 2017(EU:C:2017:592, paragraph 172), for the purposes of cross-checking PNR data against databases. See : CJEU (21 June 2022), Case C-819/19 Ligue des Droits Humains v. Conseils des Ministres; ECLI:EU:C:2022:491, § 189. On this point, AG Pitruzzella observed that the PNR Directive should be supplemented by adding those of Article 21 (1) CFR not expressly mentioned. AG Pitruzzella (27 January 2022) Opinion to Case C-817/19 Ligue des Droits Humains v Conseils des Ministres, § 231. The latter point however appears to have not been confirmed by the CJEU. The CJEU did however observe, in reference to Article 21 CFR, how Article 6(4) covers both direct and indirect discrimination: "That interpretation is, moreover, confirmed by Article 21(1) of the Charter, in the light of which the said provision must be read, which prohibits 'any' discrimination based on the said characteristics. In those circumstances, pre-determined criteria must be defined in such a way that, while worded in a neutral fashion, their application does not place persons having the protected characteristics at a particular disadvantage." CJEU (21 June 2022), Case C-819/19 Ligue des Droits Humains v. Conseils des Ministres; ECLI:EU:C:2022:491, § 197. The latter is not unimportant as PNR data might serve as proxies for sensitive information. On this point, AG Pitruzzella gives the example of air passengers' names which can "provide indications or, at the very least, allow assumptions to be made, about the ethnic origin or religious affiliation of the passenger concerned." In the same paragraph, the AG warns against the risk of stigmatisation on the basis of protected characteristics for individuals who are not suspected that data processing activities might entail. Ibid, § 184.

[813] CJEU (21 June 2022), Case C-817/19 Ligue des Droits Humains v. Conseils des Ministres; ECLI:EU:C:2022:491, § 194

[814] In this context, AG Pitruzzella does point out that the requirement of transparency clearly does not mean that the profiles should be made public. AG Pitruzzella (27 January 2022) Opinion to Case C-817/19 Ligue des Droits Humains v Conseils des Ministres, § 228.

[815] CJEU (21 June 2022), Case C-817/19 Ligue des Droits Humains v. Conseils des Ministres; ECLI:EU:C:2022:491, § 195.

[816] CJEU (21 June 2022), Case C-817/19 Ligue des Droits Humains v. Conseils des Ministres; ECLI:EU:C:2022:491, § 205.

nature of the automated processing and individual review.[817] Second, the CJEU had to consider what pre-determined criteria could be sound. The CJEU remarked how a risk assessment could take into account the *factual* conduct of persons, which following findings and expertise of the authorities involved, are indicative of a person's involvement in crime.[818] Predetermined criteria are more likely to be reliable, topical, and proportionate when they are defined taking into consideration both "incriminating" and "exonerating circumstances."[819]

One must remain cautious however in affording the status of reliability to a given piece of information. For example, in another line of cases, the CJEU had to consider the use of electronic communications traffic and location data,[820] which must also be retained for law enforcement purposes on the basis of objective and non-discriminatory criteria.[821] Here, the CJEU addressed the question whether a targeted measure for the retention of location and traffic data can be set on the basis of a geographical criterion, such as geographical areas characterised by a high risk of preparation for or commission of serious criminal offences.[822] According to the CJEU, authorities may adopt a measure on the basis of such indicators, including "the average crime rate in a geographical area, without that authority necessarily having specific indications as to the

[817] CJEU (21 June 2022), Case C-817/19 Ligue des Droits Humains v. Conseils des Ministres; ECLI:EU:C:2022:491, § 207.

[818] CJEU (21 June 2022), Case C-817/19 Ligue des Droits Humains v. Conseils des Ministres; ECLI:EU:C:2022:491, § 199-200 ; See also AG Pitruzzella (27 January 2022) Opinion to Case C-817/19 Ligue des Droits Humains v Conseils des Ministres, § 226

[819] Differentiating criteria are likely to be allowed where they concern people's factual conduct when preparing and engaging in air travel, or criteria that take into account incriminating and exonerating circumstances, as these can enhance the reliability of the information used during profiling. CJEU (21 June 2022), Case C-817/19 Ligue des Droits Humains v. Conseils des Ministres; ECLI:EU:C:2022:491., § 197 – 199.

[820] Traffic data are defined as "data processed for the purpose of the conveyance of a communication on an electronic communications network or for the billing thereof" while location data are defined as "data processed in an electronic communications network, indicating the geographic position of the terminal equipment of a user of a publicly available electronic communications service" according to Article 2(b) and (c) E-Privacy Directive, respectively. Directive 2002/58/EC of the European Parliament and of the Council of 12 July 2002 concerning the processing of personal data and the protection of privacy in the electronic communications sector (2002) OJ L201, p. 37–47.

[821] See also: CJEU (5 April 2022) Case C-140/20 G.D. v. Commissioner of An Garda Siochana; ECLI:EU:C:2022:258; CJEU (6 October 2020) La Quadrature du Net and Others v Premier Ministre and Others; ECLI:EU:C:2020:791; and CJEU (21 December 2016) Joined Cases C-203/15 and C-698/15, Tele2 Sverige AB, ECLI:EU:C:2016:970.

[822] According to the CJEU, "Those areas may include places with a high incidence of serious crime, places that are particularly vulnerable to the commission of serious criminal offences, such as places or infrastructure which regularly receive a very high volume of visitors, or strategic locations, such as airports, stations or tollbooth areas." CJEU (5 April 2022) Case C-140/20 G.D. v. Commissioner of An Garda Siochana; ECLI:EU:C:2022:258, § 79; CJEU (6 October 2020) La Quadrature du Net and Others v Premier Ministre and Others; ECLI:EU:C:2020:791, § 150.

preparation or commission, in the areas concerned, of acts of serious crime."[823] In addition, the Court claimed that "since a targeted retention using that criterion is likely to concern […] both the areas marked by a high incidence of serious crime and areas particularly vulnerable to the commission of those acts, it is, in principle, not likely moreover to give rise to discrimination, as the criterion drawn from the average rate of serious crime is entirely unconnected with any potentially discriminatory factors."[824] The latter is quite remarkable however. As Tzanou and Karyda rightfully point out, the use of geographic location as a criterion has the potential to place a disproportionate burden of surveillance and a risk of stigmatisation onto vulnerable groups, such as the poor and ethnic minorities, who often reside in so-called "crime hotspots." [825] While the use of geographic location must be non-discriminatory, "the practical consequences of such 'localized' forms of acceptable 'targeted' surveillance are most likely to be felt by the least privileged."[826]

The initial identification of reliable criteria is according to the CJEU the task of Member States.[827] Still, the CJEU has been asked to rule on the validity of certain criteria *as reliable*, and the safeguards that are put in place to ensure the reliability and non-discriminatory nature of information. In order to have a clear grasp of the impact of the situation at hand, the CJEU should have the capacity to scrutinise algorithmic decision-making systems, the logic they operate under, including their error rate, and the justification underlying their creation. If the reliability of

---

[823] CJEU (5 April 2022) Case C-140/20 G.D. v. Commissioner of An Garda Siochana; ECLI:EU:C:2022:258, § 79

[824] CJEU (5 April 2022) Case C-140/20 G.D. v. Commissioner of An Garda Siochana; ECLI:EU:C:2022:258, § 80. With regard to real-time monitoring obligations, the CJEU specified the need to have objective and non-discriminatory criteria provided for by national legislation, and subject to prior review. CJEU (6 October 2020) La Quadrature du Net and Others v Premier Ministre and Others; ECLI:EU:C:2020:791, § 189.

[825] Maria Tzanou and Spyridoula Karyda, 'Privacy International and Quadrature Du Net: One Step Forward Two Steps Back in the Data Retention Saga?', *EUROPEAN PUBLIC LAW*, n.d., 139–40. See for example: Brian Jordan Jefferson, 'Predictable Policing: Predictive Crime Mapping and Geographies of Policing and Race', *Annals of the American Association of Geographers* 108, no. 1 (2 January 2018): 1–16, https://doi.org/10.1080/24694452.2017.1293500; Rashida Richardson, Jason M. Schultz, and Kate Crawford, 'Dirty Data, Bad Predictions: How Civil Rights Violations Impact Police Data, Predictive Policing Systems, and Justice', *New York University Law Review Online* 94 (2019): 15–55. Patrick Williams and Eric Kind, 'Data-Driven Policing: The Hardwiring of Discriminatory Policing Practices across Europe.' (European Network Against Racism, 2019).

[826] Tzanou and Karyda, 'Privacy International and Quadrature Du Net: One Step Forward Two Steps Back in the Data Retention Saga?', 139–40. In reference to Maria Tzanou, 'The Future of EU Data Privacy Law: Towards a More Egalitarian Data Privacy Special Issue: Future-Mapping the Directions of EU Law: How Do We Predict the Future of EU Law?', *Journal of International and Comparative Law* 7, no. 2 (2020): 457. See also: Chapter 6, Section 6.1.5 Algorithmic Marginalisation.

[827] CJEU (5 April 2022) Case C-140/20 G.D. v. Commissioner of An Garda Siochana; ECLI:EU:C:2022:258, § 82.

information plays an important role in assessing the discriminatory nature of a decision-making process and the egalitarian harms a decision-making system might impose (over time), this type of transparency would be needed in order to make such an evaluation.[828] The CJEU's trust in geographic location as a reliable criterion for example is a reminder that the detrimental impact certain differentiation grounds can have, is not always properly understood. To make such information available would not necessarily heighten the review of the CJEU. Instead, it would provide the CJEU with access to the right, and correct amount of, information in order to make a veritable assessment and evaluation as to whether differential or similar treatment was reasonably justified, consistent, rational, and reliable.

Articulation and Proportionality

The articulation of equality's foundational underpinnings not only allows protected grounds to flourish. More generally, it could also render the proportionality approach towards equality more efficient. Where inequalities can find justification through the performance of a balancing exercise, I have argued that it is important to clearly define the exact values we want equality to achieve. Equality can only be weighed against other interests, where there is a clear understanding of the egalitarian objectives the EU wants to promote, and the specific egalitarian harms the EU hopes to guard people against. This in turn has an impact on the assessment of a measure's proportionality. The proportionality of a differentiating measure could be assessed in light of the contribution it makes to, or limitations it places upon, people's enjoyment of their egalitarian interests. Moreover, taking equality as a point of departure, the review of differential treatment could move past the differences that exist between people, to also consider the impact of, or disadvantage associated with, a given measure.

This articulation exercise might therefore facilitate the evaluation of algorithmic decision-making processes, for instance by focusing on their socio-relational or distributive impact, rather than the differentiating criteria that were used. This disadvantage-oriented approach is not entirely foreign to the EU legal system, and could find integration into a non-comparative, non-difference based proportionality review. For instance, in the PNR Directive case, the CJEU argued that the generation of differentiation grounds through fully-automated machine learning algorithms

[828] Transparency and clear accountability measures, such as impact assessments or audit trails, could provide more insight into data-driven decision-making practices, the underlying logic, and the decisional criteria used when made available. In this regard, decision-makers could also be asked to consider beforehand the delayed impact their decisions may on a socio-relational or socio-economic level. For instance, does a system risk the stigmatisation of certain groups over time? On the delayed impact of machine learning, see also: Liu et al., 'Delayed Impact of Fair Machine Learning'.

should under certain circumstances be precluded because their opacity and complexity would make it impossible for people to exercise their non-discrimination rights.[829] The egalitarian disadvantage a technology imposes onto decision-subjects, *in casu* the impossibility to exercise non-discrimination rights, was sufficient for the CJEU to conclude that their use should be forbidden. Even if a review process would remain centred on the identification and evaluation of differentiation grounds, the aforementioned articulation could be beneficial. Once the rationales that substantiate why certain grounds enjoy a higher degree of protection have been identified, the argument that people should not be differentiated on the basis of other grounds when their use would also come into conflict with those rationales can be more easily made. It does not follow from the aforementioned that the articulated egalitarian interests would emerge victorious from a balancing exercise. Instead, their role within that exercise would become more clearly defined, and hence strengthened when placed against other interests.

5.4 Conclusion

In this Chapter, I uncovered and evaluation the EU conceptualisation of equality. As a legal principle, the general principle of equality holds a primordial status within the EU framework. On the one hand, equality can be a constitutional benchmark that is invoked to protect individuals against arbitrary and irrational differentiation imposed by public authorities. On the other hand, equality can have a human-rights function that guards people against social exclusion.[830] At the same time, the EU's approach towards equality evolved over the years from equality as a tool for market integration to a more socially-oriented policy framework concerned with dignity, autonomy and respect.[831] Those functions have been incorporated into EU law - through the Charter's equality clauses, as well as in the more concretely defined, substantive EU Equality Directives – and given shape, first through a highly political decision-making process, and second through CJEU case-law. As a consequence, equality is multi-layered and complex; caught between on-going transformations and narratives which can complement and co-exist, but also conflict

[829] CJEU (21 June 2022), Case C-817/19 Ligue des Droits Humains v. Conseils des Ministres; ECLI:EU:C:2022:491, § 195.

[830] Muir, 'The Transformative Function of EU Equality Law'. Sottiaux, 'Het Gelijkheidsbeginsel : Langs Oude Paden En Nieuwe Wegen (Artikel, 2008) [WorldCat.Org]'.

[831] Prechal, 'Non-Discrimination Does Not Fall Down From Heaven: The Context and Evolution of Non-Discrimination in EU Law'.

with one another. Despite this complexity, some key guiding points can nonetheless be uncovered.[832]

THE DEFINITION OF EQUALITY AND NON-DISCRIMINATION – Unlike the ECtHR's broad approach towards equality, the EU's definition of what differentiation grounds can constitute discrimination appears closed-circuit. The CJEU has been reluctant to extend the (interpersonal) protection offered by the EU Equality Directives and the Charter (Article 20 and 21 (1) CFR) through the introduction of new (protected) differentiation grounds. Whereas the CJEU has expanded the protective scope of non-discrimination law, it has done so only in relation to characteristics the EU was expressly mandated to govern. This hesitance can be motivated. Among others, it could breach the established boundaries of EU law, as well as disrupt the balance the EU wants to maintain between market-oriented objectives on the one hand, and social goals on the other hand. Unfortunately, the latter motivations cannot explain why the CJEU thus far fails to recognise, within the bounds of its competences, more complex forms of discrimination, such as intersectional discrimination based upon (or linked to) protected characteristics.

In cases where equality functions as an institutional check, it could have a more open-ended and sweeping scope, and therefore hold an evaluative function in relation to non-personal, status-based differentiation grounds generated or imposed within (public) decision-making contexts. Yet the potential of this function is not entirely clear. If the goal is to guard people against the egalitarian harms imposed via algorithms, this benchmarking function would converge with a more socially oriented conceptualisation of equality. Given the reluctance to expand the protective scope of equality to new status-based grounds in the latter domain, it is uncertain to what extent protection would be extended to non-status-based grounds in the former. At the same time, case-law in the field of data-driven domains illustrate the CJEU's willingness to address the (non-)discriminatory nature of non-personal and non-protected differentiation grounds and their impact on people, such as the use of geographical location in the field of law enforcement. Hence, it is likely that the CJEU will be asked to do this more in the future.

THE JUSTIFICATION OF INEQUALITY AND NON-DISCRIMINATION – The EU's conceptualisation of equality offers protection against various socio-relational egalitarian harms (e.g., social exclusion, prejudice, stereotypes, stigma and vulnerability) people might experience by virtue of, or in connection to their possession of suspect or protected characteristics, and in relation to the

[832] As aforementioned, this conceptualization does not consider the difficulty that may exist in applying equality and non-discrimination concepts onto algorithmically guided decision-making systems.

distribution of a set of well-defined prized or justice-relevant goods.[833] Through these mechanisms, EU equality laws look to preserve individuals' moral worth and their capacity to exercise autonomous choice by reducing their risk of being placed in a position of social and distributive inequality in their interpersonal relations.[834] In light of the risk that algorithms present as producers of egalitarian harms, whereby these harms can be imposed on a variety of differentiation grounds and in a wide range of decision-making domains this constrained approach might not be sufficiently robust.

The level of protection offered could be extended through Articles 20 and 21 CFR, which enable the CJEU to interpret the general principle of equality as an institutional or interpersonal benchmark in domains governed by EU law. Here, equality could be further defined through a balancing of interests: the differential treatment of comparable situations (or vice versa) can be justified where a reasonable and objective justification can be provided. Whereas differentiation on the basis of protected characteristics could remain subject to a heightened level of scrutiny, in the case the CJEU would be in the possibility to evaluate (non-suspect) algorithmically generated differentiation grounds and associated inequalities. I argue that a low-level, rationality-based approach would be ill-suited. Instead, a genuine balancing of interests ought to be preferred, where the differentiation ground in question plays a (lesser) role and more attention can be paid to the egalitarian harms a system produces.[835] In this context, the CJEU is correct illustrates the importance of institutional safeguards: algorithmically guided decisions should be made transparent and traceable in order to assess among others the reliability, soundness and desirability of the differentiation grounds used (including to what extent a given piece of information can contribute to the goal achieved), the egalitarian harms a system can produce on a socio-relational and distributional level, the system's potential to interfere with people's capacity to exercise their non-discrimination rights, and the correspondence of that system's design and outcomes with the (competing) interests and goals pursued by a decision-maker.

---

[833] Where a differentiation ground and sector are covered by a EU Equality Directive, there are two possibilities. In a case a decision-making act is directly discrimination, inequality cannot be justified, unless an express exemption applies. In the case of indirect discrimination, a proportionality test will follow. With regard to the heightened protection offered to the protected characteristics, the analysis showed that their protection is founded upon a plurality of views similar to those that underpin the ECtHR's conceptualisation of equality.

[834] A question that has not been explored is to what extent the symmetrical protection offered by protection grounds might come into conflict with substantive equality goals that look to protect historically disadvantaged groups. Similarly, this Chapter has not addressed the possibility of positive action measures, which the EU Equality Directives allow Member States to introduce, and their relationship with algorithmic decision-making systems.

[835] On the advantages of an open-ended proportionality test for the evaluation of algorithmically generated harms, see also: "Chapter 4, Section 4.4 Conclusion."

PART III – Navigating Political Philosophy: Socio-Relational and Distributive Conceptualisations of Equality and the Articulation of Algorithmic Egalitarian Harm.

Introduction

In **Part I**, I illustrated how algorithmic decision-making systems may obstruct decision-subjects in their enjoyment of certain fundamental moral, socio-relational and/or distributive egalitarian interests. In **Part II**, I evaluated the legal interpretation of equality and non-discrimination and demonstrated how it may not be sufficiently robust to address the egalitarian harms and disadvantage algorithmic decision-making systems risk to impose onto a wider range of decision-subjects. In **Part III,** I investigate and articulate how those harms may arise in practice. For this, I once again turn to my multidimensional concept of equality, which functions here as a navigational beacon to locate specific political philosophical conceptualisations of equality that can further deepen our understanding of the egalitarian harms algorithms risk to impose. Focusing on specific socio-relational and distributive interests decision-subjects may share, I demonstrate how algorithmic decision-makers, as a result of the choices they make and the social changes they initiate, can interfere with these interests. Knowing what actions and behaviours obstruct equality and why, we can then formulate recommendations to both decision-makers and regulators regarding the responsibilities they have to strengthen and empower decision-subjects in their confrontations with the algorithmic environment.

First, I explore the socio-relational egalitarian harm algorithmic decision-makers may impose through an exploration of the notions of domination and oppression, including (unjustified) deindividualisation and misrepresentation **(Chapter 6).** I then examine how algorithms in their function as producers of opportunities, benefits and burdens may interfere with people's access to, or enjoyment of justice-relevant goods via the distributive ideals of "equality of primary goods" and "equality of capability" **(Chapter 7).**[836]

These specific notions of equality can take a twofold function. On the one hand, each distinct conceptualisation provides an alternative account to the law. Like the normative underpinnings underlying European equality and non-discrimination law, they can be used to articulate the egalitarian harms decision-subjects risk to incur within the algorithmic environment, the actions

[836] While the focus will be on socio-relational and distributive equality, moral equality keeps its patrolling and review function. Once a harm is identified as problematic, moral equality looks to guarantee that no person is (unjustifiably) excluded from being protected against that harm. Similarly, algorithmic decision-makers in their obligation to accommodate, rather than to interfere into people's fundamental interests, should not unjustifiably disfavour or arbitrarily exclude people in their enjoyment of these interests.

and behaviours that prevent equality coming to fruition, and the institutional safeguards and resources decision-subjects need access to in order to strengthen their position vis-à-vis decision-makers.[837] On the other hand, since deficiencies in the legal approach to equality have been uncovered, these notions can also inform, complement or expand upon the law. As European equality law recognises both the socio-relational and distributive dimensions of equality, examining these conceptualisations of equality within political philosophy can lead to a more holistic and robust response towards the regulation of algorithmic decision-making

In this context, it may be noted how the conceptualisations described here appear to have at least one advantage over the law in light of its deficiencies. **In Part II,** I discovered how the language of the law frames equality as a notion that is concerned with the protection of (more or less clearly defined) salient social groups.[838] As a consequence, the legal interpretation of equality and non-discrimination may not be sufficiently robust to address the egalitarian harms and disadvantage algorithmic decision-making systems risk to impose onto a wider range of decision-subjects - whether they are (members of) more complexly defined socially disadvantaged groups or non-representational, intangible or seemingly random algorithmic aggregates. Furthermore, taking into consideration the various domains in which algorithms can find use, the scope *ratione materiae* of European equality law may be too narrowly construed. This appears even more problematic given the extended impact decision-making systems may have due to the interconnected nature of the algorithmic environment: a decision-making system may introduce inequalities in one domain, but their actual social impact can take place in another. The protective scope offered by the conceptualisations analysed in **Part III** is not limited by, nor based upon, a defined list of (identifiable) specific characteristics or traits. In addition, their application is not (directly) limited to a finite set of sectors. Instead, their protective scope is defined by the underlying fundamental interest that must be equalised and the specific disadvantage people might experience when this equalisation is not maintained. Hence, any algorithmic operation that interferes with these interests, regardless of the ground used, group captured, or domain in which the algorithm applies, represents a risk of harm to those involved, and they would therefore have a rightful claim to have this interference made subject to scrutiny. Because the egalitarian harms they want to guard people against can originate, and take effect, in a wide variety of domains., their monitoring function is more likely to transcend the protective borders of equality and non-discrimination laws.

---

[837] As these Chapters will demonstrate however, it is possible to discern various synergies between these conceptualisations. Where this is the case, their complementary nature will be highlighted.

[838] See on this point: Wachter, 'The Theory of Artificial Immutability'.

6 Socio-Relational Conceptualisations of Equality and Algorithmically Guided Decision-Making

The ideal of social equality reflects the belief that a just society is comprised of free and equal persons and every person is respected, recognised and valued as a member of that society.[839] Following a socio-relational conceptualisation of equality, the rules that govern society must be such that they enable people to maintain their status as social equals within the various relationships they hold.[840] On the one hand, we should define the (positive) institutions, principles, conditions, practices and rights people need to enjoy, or have access to, in order to stand in equal social relations.[841] On the other hand, we should strive to abolish (negative) actions, treatment, behaviours, structures and phenomena that enable individuals, (social) groups, private corporations or governments to threaten or undermine a given person or group's equal social standing.[842]

Framed as a negative ideal, socio-relational conceptualisations of equality highlight the conditions under which algorithmic decision-makers (risk to) erode the equal social standing of decision-subjects. Such egalitarian harms can arise from the way algorithmic decision-making structures are imposed onto decision-subjects, as well as the particular outcomes these structures generate for them. Framed as a positive ideal, equality can be achieved by giving decision-subjects access to a set of social and institutional resources that can strengthen their position as social equals within an algorithmically mediated environment. These conditions enable decision-subjects to

---

[839] Schuppert, 'Non-Domination, Non-Alienation and Social Equality', 444.

[840] These harms not only take effect in people's interaction with others, e.g., a stereotype, but may also have an impact on a person's relationship with themselves, e.g., a loss of self-respect. Scheffler, 'The Practice of Equality', 21; Scheffler, 'What Is Egalitarianism?', 2003. Garrau and Laborde, 'Relational Equality, Non-Domination, and Vulnerability', 48. Baker, 'Conceptions and Dimensions of Social Equality', 77. See also: Fabian Schuppert, 'Suffering from Social Inequality: Normative Implications of Empirical Research on the Effects of Inequality', *Philosophical Topics* 40, no. 1 (2012): 97–115. Schuppert, 'Non-Domination, Non-Alienation and Social Equality'; in reference to Martin O'Neill, 'What Should Egalitarians Believe?', *Philosophy & Public Affairs* 36, no. 2 (March 2008): 119–56, https://doi.org/10.1111/j.1088-4963.2008.00130.x.

[841] Anderson, 'What Is the Point of Equality?', 313; Scheffler, 'The Practice of Equality', 21–22.

[842] Young was among the first to argue that social justice should be concerned with structures of oppression and domination. Anderson later identified oppression and domination as enemies of social equality. Phrased positively, social egalitarians try to create a community where people stand in equal relationships towards one another by formulating principles conducive to achieving this goal. Anderson, 'What Is the Point of Equality?', 313. Young, *Justice and the Politics of Difference*. Anderson, 'What Is the Point of Equality?', 313.

uphold, or regain, their status as social equals in all their interactions, and not only in their confrontation with more powerful decision-makers. In sum, these notions inform the evaluation and regulation of the various relationships decision-subjects hold within the algorithmic environment: the vertical relationships they enter into with decision-makers and the relationships they hold with other decision-subjects as part of, and within, individual decision-making processes and larger algorithmic structures.[843]

In the following Sections, I examine and evaluate algorithmically guided decision-making structures and processes by focusing on the socio-relational harms of oppression and domination, as well as deindividualisation and misrepresentation.[844] Equality is defined as a duty to equalise decision-subjects' fundamental interest to be free from algorithmically (induced) forms of oppression, domination, deindividualisation and misrepresentation. First, this Chapter clarifies the notions of algorithmic oppression and domination (Section 6.1). I explain what algorithmic acts or structures of oppression and domination are, why these acts ought to be considered socio-relational status harms (as well as a type of vulnerability), and how such harms can be inimical to decision-subjects' fundamental interests. I then specify more concretely how algorithmic decision-makers can oppress and dominate. Following this analysis, I propose safeguards that decision-makers and regulators could consider protecting decision-subjects from suffering socio-relational harms, while at the same time, bolster decision-subjects' ability to regain control over their living environment. Second, I investigate the socio-relational status harms algorithms risk to impose onto decision-subjects due to their generalising nature (Section 6.2). I investigate the relationship that exists between stereotypes and generalisations and explain whether and how the latter might result in socio-relational status harms that are similar to those associated with the former. I articulate the negative impact of generalisations through the notions of de-individualisation and misrepresentation. Finally, I consider to what extent the safeguards identified within this Chapter's first section could also benefit decision-subjects against these harms.

---

[843] See also Chapter 2, Section 2.2.1 "Algorithmic Socio-Relational Harms." See also: Viljoen, 'A Relational Theory of Data Governance'.

[844] These notions are not only inimical to the relationships people hold, but they can also lead to socio-economic inequalities. In principle, social equality as a positive ideal could achieved through distributive theories of justice, e.g., by making sure everyone has the same opportunities, or through democratic processes of inclusion and participation. While equality positively phrased could certainly inform algorithmic decision-making, it is more interesting to see how algorithms could act in a manner that is hostile to social equality, and take this as a point of departure, and from thereon out see how these notions can inform algorithms in a positive manner.

6.1 Algorithmic Oppression and Domination[845]

Algorithmic oppression can take multiple forms. The notion can refer to the capacity of decision-makers to i) exploit, ii) marginalise, iii) render powerless decision-subjects (or take advantage of their powerlessness), iv) impose onto them their (subjective) world views in ways that render decision-subjects' lived experiences invisible (cultural imperialism), or v) expose decision-subjects to (digital) forms of violence. Acts and structures of oppression constrain people in their right towards self-development, understood as a person's capacity to learn, develop, exercise, and express their skills, capacities, and experiences in socially meaningful and recognised ways.[846] The power to oppress typically implies a power to dominate.[847]

Algorithmic domination can be understood as the capacity for decision-makers to interfere in a person's valuable options and choices in a manner that is arbitrary, unchecked or uncontrolled.[848]

---

[845] This section draws from the work of Philip Pettit (Non-Domination) and Iris Marion Young (Non-domination and non-oppression) in order to define and conceptualise equality understood as the capacity for people to be free from domination and oppression. The notions of self-development and self-determination are taken from Young.

[846] These are what Young refers to as the five faces of oppression: Young, *Justice and the Politics of Difference*, chap. 2. According to Young. oppression refers to "systemic institutional processes that prevent some people from learning and using satisfying and expansive skills in socially recognized settings, or institutionalized social processes which inhibit people's ability to play and communicate with others or to express their feelings and perspective on social life in contexts where others can listen." Young, 37. In reference to Carol C. Gould, *Rethinking Democracy: Freedom and Social Cooperation in Politics, Economy, and Society* (Cambridge University press, 1988); William Arthur Galston, *Justice and the Human Good* (University of Chicago Press, 1980).

[847] Young, *Justice and the Politics of Difference*, 38. Most people will have to enter into hierarchical relationships of power and domination at one point or another. While these relationships can function as a constraint towards people's right to self-determination, they do not necessarily render a person voiceless or invisible. Oppressed groups however are voiceless and invisible, and as such, more vulnerable to being dominated.

[848] Domination is a social, structural, and institutional process as it comes about through the actions, whether intended or unintended of many people. According to Young, People live within structures of domination when other persons or groups have the capacity to "determine without reciprocation the conditions of their actions, either directly or by virtue of the structural consequences of their actions." On a surface level, this definition of domination corresponds to Pettit's definition of domination: "someone, A, will be dominated in a certain choice by another agent or agency, B, to the extent that B has a power of interfering in the choice that is not itself controlled by A." Young and Pettit differ however in how they view non-domination. Pettit understands domination primarily as an agential, rather than a structural phenomenon. Moreover, he considers freedom from non-domination to be a personal and primary good which everyone should value. Young rejects the idea that non-material objects like power or "freedom of non-domination" should be seen as distributable goods or things. Doing so would obscure their social and relational character. Young, 31–32 and 38. In reference to Wartenburg. Philip Pettit, *On The People's Terms: A Republican Theory and Model of Democracy* (Cambridge, New York, Melbourne, Madrid, Cape Town, Singapore, Sao Paolo, Delhi, Mexico City: Cambridge University

Domination is a power decision-makers hold over decision-subjects to actively interfere with their right to self-determination, which can be understood as a person or group's capacity to exercise control over, and participate in determining, their own options, choices and the conditions of their actions.[849]

Algorithms can be tools for domination and oppression. This concern not only applies to isolated acts or decisions. Domination and oppression are structural or institutional phenomena.[850] This understanding recognises how algorithmic systems are embedded within existing economic, social, and political cultures, from which they take meaning. In order to attain algorithmic justice decision-makers should avoid the replication and reinforcement of existing structures of oppression and domination experienced by (historically) disadvantaged social groups. Those injustices arise as a result of the many (un)conscious actions, attitudes, assumptions or reactions of people and institutions as part of everyday life.[851] The algorithmic environment should not become yet another layer where entrenched inequalities are reproduced.[852] Structural inequalities, such as (distributive) inequalities in wealth, social and economic power, but also inequalities in access to knowledge (like those present within the algorithmic environment) are unjust as they can help to introduce, maintain or reinforce oppression.[853] Like injustice comes about through the support and actions of many third parties, so does the institutionalised power that algorithmic decision-makers have to oppress and dominate others. For instance, if data protection laws are not enforced, decision-makers are given the opportunity to harvest and exploit the personal data of their decision-subjects. At the same time, decision-subjects often willingly enter into data-driven relationships even when they know the decision-makers might not have their (best) interests in mind.

Press, 2012), 50. Philip Pettit, *Republicanism: A Theory of Freedom and Government*, *Republicanism* (Oxford University Press, 1999), 52, https://www.oxfordscholarship.com/view/10.1093/0198296428.001.0001/acprof-9780198296423. Young, *Inclusion and Democracy*, chap. Democracy and Justice.

[849] Young, *Justice and the Politics of Difference*, 37.

[850] Young, chap. 1.See also: Young, 'Equality of Whom? Social Groups and Judgments of Injustice', 13.

[851] Young, *Justice and the Politics of Difference*, 41.

[852] Historically disadvantaged and social groups experience the greatest risk of domination and oppression. In the following sections, I nonetheless hope to illustrate how the algorithmic environment can set the conditions for non-social groups, such as intangible algorithmic aggregates, to become subject to (at least singular acts of) domination and oppression.

[853] Young, *Inclusion and Democracy*, 34. See also: Young, 'Equality of Whom? Social Groups and Judgments of Injustice'. This is a goal most distributive justice and non-discrimination theorists would subscribe too.

6.1.1 Oppression and Domination: Vertical and Horizontal Phenomena

VERTICAL DIMENSION - Vertical algorithmic oppression and domination concern situations where decision-makers directly oppress or exert dominance over decision-subjects. On the one hand, the algorithmic environment may simply replicate a type of domination or oppression that already existed before. An employer might have held a dominating power to change the work schedule of its employees unilaterally and arbitrarily. In a digitally mediated environment, they can now outsource this capacity to an algorithmically guided time management system.[854] On the other hand, the datafication of society has given rise to asymmetries in power which have placed a new range of decision-makers in a position where they can directly interfere with people's ability to exercise self-determination or self-development. Social media platforms have the power to mediate people's relationships and participation with others. A decision-maker's ability to do so signifies the existence of an inegalitarian social relationship between the decision-maker and decision-subjects.

Due to their unique properties, algorithms have expanded the ability of decision-makers to oppress and dominate others to their own and others' ends. Gig platforms can exploit their workforce more efficiently by controlling the strategic choices workers would make.[855] Dynamic pricing mechanisms within app-based delivery and hail ride services can be purposefully designed to keep workers active at specific hours in order to maximise economic returns.[856] Inferential analytics can be used to obtain (sensitive) personal information regarding decision-subjects in order to sell this information for (behavioural) advertising purposes. An oppressive regime might employ data mining technologies to scrape social media platforms in an attempt to uncover the presence, activities, and whereabouts of political dissidents.

HORIZONTAL OPPRESSION AND DOMINATION – Through their power to establish the conditions under which people are differentiated against, algorithmic decision-makers can also cultivate positions

[854] See also: Eike Gräf, 'When Automated Profiling Threatens Our Freedom':, *European Data Protection Law Review* 3, no. 4 (2017): 446, https://doi.org/10.21552/edpl/2017/4/6.

[855] James Muldoon and Paul Raekstad, 'Algorithmic Domination in the Gig Economy', *European Journal of Political Theory*, 23 February 2022, 14748851221082078, https://doi.org/10.1177/14748851221082078.

[856] Muldoon and Raekstad.

or structures of domination or oppression within the broader social environment.[857] Algorithmic decision-makers that control the distribution of opportunities, burdens and benefits among people effectively govern the resources and means people need access to in order to exercise meaningful choice in life. Such decisions also have an impact (within a given context) on people's ability to meaningfully participate within society. Suppose a highly popular credit scoring algorithm uncovers on the basis of some newly discovered graphological or stylometric technique that left-handed people are less trustworthy.[858] Left-handed people find themselves unable to get a loan. This inability would reduce their capacity to exercise choice and participate in social life. Put differently, algorithmic decision-makers can reinforce or restructure existing power dynamics within society as a result of their choices and actions. This means they also have the capacity to render certain people and groups more vulnerable than others, and consequently, more prone towards domination and oppression, whether they intend to or not.[859]

6.1.2 A Function of Power?

Positions of domination and oppression are formed through a totality of social institutions, relationships, and social processes within which people operate and navigate. Due to their ubiquity, algorithmic decision-making structures have become an integral component of this totality. Against this background environment, a decision-maker's capacity to dominate or oppress, or contribute to domination or oppression, is a function of the power they hold. This power can be defined in terms of decision-makers' access to the infrastructures, decisions and knowledge needed to engage into automation.[860] As aforementioned however, these positions of power can be acquired and maintained through the actions of many third parties.[861] They will

---

[857] Antoinette Rouvroy and Yves Poullet, 'The Right to Informational Self-Determination and the Value of Self-Development: Reassessing the Importance of Privacy for Democracy', in *Reinventing Data Protection* (Dordrecht: Springer, 2009), 45–76.

[858] See for instance: Anastasia Siapka, 'Show Me How You Write and I'll Tell You Who You Are', CITIP blog, 24 February 2022, https://www.law.kuleuven.be/citip/blog/show-me-how-you-write-and-ill-tell-you-who-you-are/.

[859] An algorithmic decision-maker might be more concerned with the efficiency of a given decision-making system, or the economic gains a system can produce, rather than the vulnerability-status they inflict onto others. Yet, their impact on the latter still remains tangible. Hence, proponents of social equality actively promote ideals of participation and inclusion as a part of an algorithm's design. By incorporating the views of those who will be affected by a given decision-making process, decision-makers can better gauge how they can impact people's right to self-determination and development.

[860] See Chapter 1, Section 1.2.3 "Algorithmic Power: Infrastructures, Decisional Control and Knowledge Creation."

[861] See above, Section 6.1 Algorithmic Oppression and Domination

contribute to the political importance, legal protection, financial means, social presence algorithmic decision-makers are afforded.

The asymmetries in power that make it possible for decision-makers to oppress and dominate others can be equalised through strong (and enforceable) regulatory frameworks. Gig-workers for instance have had little recourse against gig-platforms because their precarious status and working conditions are often left unacknowledged under current labour laws. In certain cases, gig-workers did find success through the exercise of data subject rights granted to them under EU data protection laws, such as the right to access and the right not to be subject to automated decision-making practices.[862] These cases illustrate how the law can play a crucial role in breaking down power asymmetries and informational barriers. Similarly, equality and non-discrimination laws could become an invaluable tool in combatting the egalitarian harms decision-subjects can experience through algorithmic acts of domination and oppression. Where the legal environment allows algorithmic decision-makers to acquire unbridled power, it inevitably places their decision-subjects in inegalitarian social relationships. Furthermore, the greater their power, the greater their restructuring capacity, and the higher the risk that the differentiation they impose onto decision-subjects translates into concrete socio-relational inequality on the horizontal plane.

Second, the power to oppress and dominate can be linked to the control a given decision-maker has over people's fundamental interests. Where a given society conceptualises equality as a policy ideal that guards people against acts and structures of domination and oppression, the fundamental interest society aims to equalise among its citizens is their capacity towards self-

---

[862] See for example: Laurens Naudts, Pierre Dewitte, and Jef Ausloos, 'Meaningful Transparency through Data Rights: A Multidimensional Analysis', in *Research Handbook on EU Data Protection Law* (Edward Elgar Publishing Limited, 2022), 530–71. Toh, 'Empowering Workers Through Digital Rights', *Digital Freedom Fund* (blog), accessed 27 September 2022, https://digitalfreedomfund.org/empowering-workers-through-digital-rights/. See: Rb. Amsterdam 11 March 2021, ECLI:NL:RBAMS:2021:1018; <https://uitspraken.rechtspraak.nl/inziendocument?id=ECLI:NL:RBAMS:2021:1018> accessed 23 June 2021.; Rb. Amsterdam 11 March 2021, ECLI:NL:RBAMS:2021:1019; <https://uitspraken.rechtspraak.nl/inziendocument?id=ECLI:NL:RBAMS:2021:1019> accessed 23 June 2021; Rb. Amsterdam 11 March 2021, ECLI:NL:RBAMS:2021:1020; < https://uitspraken.rechtspraak.nl/inziendocument?id=ECLI:NL:RBAMS:2021:1020> accessed 23 June 2021. For more information, including an unofficial translation of the aforementioned cases into English, see also: Ekker Legal, 'Dutch court rules on data transparency for Uber and Ola drivers.' (Ekker Legal, 13 March 2021) <https://ekker.legal/2021/03/13/dutch-court-rules-on-data-transparency-for-uber-and-ola-drivers/> accessed 23 June 2021. See also: App Drivers & Couriers Union, 'Gig Economy Workers score historic digital rights victory against Uber and Ola Cabs.' (12 March 2021), https://www.adcu.org.uk/news-posts/gig-economy-workers-score-historic-digital-rights-victory-against-uber-and-ola-cabs accessed 23 June 2021

determination and self-development. Evidently, the power to oppress and dominate gives decision-makers direct control over the way decision-subjects can exercise their right to self-determination and self-development. However, power might also be defined in relation to the control a decision-maker, in their function as a distributor of burdens, benefits and opportunities, holds over the justice-relevant goods that people need access to in order to enjoy or realise these fundamental interests.[863]

6.1.3 Socio-Relational Status Harms: A Type of Vulnerability

Acts and structures of domination and oppression signal the existence of unequal social relationships, i.e., hierarchies within society that allow a person or group to become subject to the (unilateral and uncontrolled) power of others to the detriment of their fundamental interest towards self-development and self-determination. In a society of equals, these fundamental interests ought to be enjoyed by all with equal intensity.

A person or group that is likely to become the subject of domination or oppression suffers a status-inequality. They carry a vulnerability-marker that indicates their unequal social standing.[864] People who find themselves free from domination and oppression are non-vulnerable: they are recognised and respected as being of equal (or higher) social standing, and therefore less likely have their fundamental interests interfered with. They are neither voiceless, nor invisible; they

[863] In this regard, reference can be made to the distributional theories of Rawls and Nussbaum discussed in the following Chapter. Both recognise the importance of people's self-development and self-determination. Whereas Rawls protects the aforementioned rights through the distribution of a set of primary goods, Nussbaum argues that people should be able to enjoy a set of key substantive freedoms. Hence, where algorithms exert control over primary goods or central capabilities, the risk that the inequalities they generate are more likely to contribute to a given person's social inequality increases. Regarding the latter point, it is nonetheless important to take into consideration Young's critique of the distributive paradigm, which she believed to presuppose and obscure the importance that social structures, relationships, institutional contexts, decision-making contexts, and culture hold. In addition, Young argued that the distributional logic is often wrongfully applied onto non-material goods like self-respect and power. These phenomena are better explained as a function of rules, relationships and social processes, rather than access to particular goods. Young, *Justice and the Politics of Difference*, 25. In reference to Hartstock, 1983.

[864] The notion vulnerability is particularly present in the work of Pettit. Pettit has developed the notion in relation to non-domination, self-determination, and autonomy, one can understand this notion also in relation to oppression. Garrau and Laborde however recognise how vulnerability may also relate to oppression. Garrau and Laborde, 'Relational Equality, Non-Domination, and Vulnerability'.

can meaningfully participate to social life, and exercise choice and control over their actions.[865] In contrast, persons and groups that experience oppression and domination are vulnerable. Their actions and behaviour are curtailed. They face exploitation, marginalisation, unjustified stereotypes and stigma, or violence. [866] Concerning domination, the (possibility) for others to arbitrarily interfere in a person or group's choices, imposes uncertainty, which can cause people to act more cautious, or subservient to the actions of those in power. As a result of these confrontations, people are made voiceless and invisible in their navigation of social life: they are dismissed and denied the opportunity to be recognised, heard and worth listening to. [867] Domination, oppression, and vulnerability share a reciprocal relationship. Whereas persons or groups who are in a (socially) vulnerable position are more likely to become the subject of domination and oppression, acts and structures of domination and oppression in turn render them more vulnerable.[868]

---

[865] See for example: Pettit, *Republicanism*, 82–89. See also: Garrau and Laborde, 'Relational Equality, Non-Domination, and Vulnerability'; Schuppert, 'Non-Domination, Non-Alienation and Social Equality'; Fabian Schuppert, *Freedom, Recognition and Non-Domination*, vol. 12, Studies in Global Justice (Dordrecht: Springer Netherlands, 2014), https://doi.org/10.1007/978-94-007-6806-2. Cécile Laborde, 'Political Philosophy, Social Theory, and Critical Republicanism', in *Critical Republicanism* (Oxford: Oxford University Press, 2008), https://doi.org/10.1093/acprof:oso/9780199550210.003.0001; Schuppert, *Freedom, Recognition and Non-Domination*, 12:443. Pettit, *Republicanism*, 91. On recognition, see also the work of Honneth and Fraser. Honneth, 'Recognition and Justice'; Marcus Ohlström, Marco Solinas, and Olivier Voirol, 'Nancy Fraser and Axel Honneth's Redistribution or Recognition? A Political-Philosophical Exchange1', *Iris* 3, no. 5 (2011): 205–21; Fraser, 'Social Justice in the Age of Identity Politics. Redistribution, Recognition, Participation.'; Nancy Fraser, 'Recognition without Ethics?', *Theory, Culture & Society* 18, no. 2–3 (1 June 2001): 21–42, https://doi.org/10.1177/02632760122051760.

[866] Young, 'Equality of Whom? Social Groups and Judgments of Injustice', 9–11. In reference to: Jean Hampton, *Political Philosophy*, vol. 108 (Westview Press, 1997); Marilyn Frye, *The Politics of Reality: Essays in Feminist Theory* (The Crossing Press, 1983). On psychological harms, see: Ann E. Cudd, 'Psychological Harms of Oppression', in *Analyzing Oppression* (New York: Oxford University Press, 2006), https://doi.org/10.1093/0195187431.003.0006; Ann E. Cudd, *Oppression: The Fundamental Injustice of Social Institutions*, *Analyzing Oppression* (Oxford University Press), accessed 7 June 2021, http://oxford.universitypressscholarship.com/view/10.1093/0195187431.001.0001/acprof-9780195187434-chapter-1.

[867] Conversely, a person who is dominated will have a much harder time in having their claims taken serious. Pettit, *Republicanism*, 91. Schuppert, 'Non-Domination, Non-Alienation and Social Equality', 445. On recognition, see also the work of Honneth and Fraser. Honneth, 'Recognition and Justice'; Ohlström, Solinas, and Voirol, 'Nancy Fraser and Axel Honneth's Redistribution or Recognition?'; Fraser, 'Social Justice in the Age of Identity Politics. Redistribution, Recognition, Participation.'; Fraser, 'Recognition without Ethics?'

[868] Pettit makes an interesting observation regarding a person's vulnerability status, their being dominated and their being historically subject to oppression. A person will not be able to enjoy non-domination if the people that belong to the same "vulnerability class" do not enjoy non-domination either: "to the extent that those others are exposed to arbitrary interference, you too are exposed; to the extent that they are dominated, you too are dominated." Certain (socially salient) groups carry a so-called "badge of vulnerability." While not every member may be dominated in

Domination and oppression are also relational phenomena in the sense that they are experienced in relation to others who are more privileged. [869] Conversely, policies that look to ensure freedom from oppression and domination should investigate the power dynamics that characterise a given environment and establish those conditions and practices necessary to reduce the vulnerabilities that have emerged from these power imbalances.[870]

6.1.4 Equality as Freedom from Algorithmic Domination and Oppression?

In order to contribute to people's right to self-determination and self-development, algorithmic justice requires that the algorithmic environment is structured in such a way that it is free from domination and oppression: people should be equal in their enjoyment of non-oppression and non-domination.[871] Hence, algorithmic decision-makers are called upon to consider the consequences of their actions: to what extent, do my decisions oppress or dominate?

On the one hand, algorithmic decision-makers should not exercise their powers in a way that dominates or oppresses their decision-subjects. Direct acts of algorithmic oppression and domination are unjustified not only because they directly interfere with people's fundamental interests, but they are also unjustified because the algorithmic decision-maker, whether it be a government or private entity, takes advantage of their privileged socio-relational position, and the vulnerability of their decision-subjects. In this context, Viljoen also notes how "by forming and

practice, this badge signifies that a given person can be dominated. Non-domination must thus be seen as a common good as it cannot be achieved individually. Here, Pettit does draw from Young's work on the oppression of social groups: *"The goal of freedom as non-domination gives a common cause to each of the salient vulnerability classes in any contemporary society: to each of those groups who are rendered particularly vulnerable […] to each of those groups who […] count as relatively oppressed."* Pettit, *Republicanism*, 122. Young, *Justice and the Politics of Difference*, chap. 2. Pettit, *Republicanism*, 122–24.

[869] See also: Young, *Justice and the Politics of Difference*, 42.

[870] For instance, Pettit believes that by promoting freedom as non-domination, the political significance of other "status" markers linked to historically disadvantaged groups, like class, gender, ethnicity should decrease. Pettit, *Republicanism*, 125. See also: Pettit, *On The People's Terms: A Republican Theory and Model of Democracy*, 63–90.

[871] For a more specific analysis on the oppression of algorithms through racism and sexism, see also: Noble, *Algorithms of Oppression*.

then acting on population-level similarities in oppressive and dominating ways, datafication may materialize classificatory acts of oppressive-category formation that are themselves unjust."[872]

On the other hand, decision-makers should be mindful of the powers they hold. They should make sure that their actions or decisions do not intensify, or introduce, the risk for a person or group to become the (future) subject of domination and oppression. Particular attention must be paid to differentiating acts that rely upon the placement of people into population-based or group-level relationships as in doing so they directly interfere with the relational fabric of a given society. Depending upon the context in and the conditions under which a decision is made, these interventions can maintain, decrease, or increase a person or group's vulnerability status vis-à-vis others. In this context, algorithmic differentiation acts, and the inequalities they produce, are more likely to constitute an unjustified harm the more they increase the vulnerability status of a person or group. Given this reality, Viljoen observes that decision-subjects might have a "collective interest against the unjust social processes data flows may materialize, against being drafted into the project of one another's oppression as a condition of digital life, and against being put into data relations that constitute instances of domination and oppression for themselves or others on the basis of group membership."[873]

6.1.5 Algorithmic Oppression[874]

Acts and structures of oppression constrain persons in their right towards self-development. Drawing from Young's definition of oppression as referring to a family of concepts and conditions - the five faces of oppression: exploitation, marginalisation, cultural imperialism, powerlessness, and violence – in this Section, I examine how algorithmic decision-making systems can lead to oppression. According to Young, oppression immobilises social groups, and each face of oppression can be operationalised as a means to assess whether a social group faces (digital) oppression. I furthermore illustrate how similar harms could also be imposed onto non-tangible or non-representational aggregates that do not (necessarily) share a link with prior injustice. Whereas these harms might not constitute oppression *sensu stricto*, they could be referred to as *oppression-like*.[875] Each notion can function as a lens through which decision-makers reflect upon

[872] Viljoen, 'A Relational Theory of Data Governance', 643.

[873] Viljoen, 651.

[874] This section builds upon the work of Iris Marion Young, *Justice and the Politics of Difference*, chap. 2.

[875] Young, 64. Being subject to any of the aforementioned forms of oppression is in principle sufficient for a group to be considered oppressed.

their powers and responsibilities prior to the design and deployment of a decision-making system. After deployment, they can be used to monitor and review the socio-relational impact of decision-making systems.

Exploitation of the Powerless

EXPLOITATION - Exploitation is a process through which the efforts or energy of one person or group is transferred to enlarge the power of others on a social and economic level. It concerns the situation where some people find that they exercise "their capacities under the control, according to the purposes and for the benefit of other people."[876] Through the various (design) choices that feed into the creation and subsequent deployment of an algorithmic decision-making system, decision-makers exercise control over their decision-subjects, (help) define the purposes for which this control is established, and who this control will ultimately benefit.

There are a variety of ways through which algorithmic decision-making systems can be exploitative.[877] The data underlying a system can be collected in a manner that was exploitative for instance. Suppose an eHealth start-up is about to have a breakthrough in their development of a revolutionary deep learning algorithm that can detect emerging health conditions at an early stage. In order to continue their research however, they need access to a lot of DNA data. Due to their unproven history, they have little success in convincing volunteers to have their health data sampled. Not wanting to give up, the start-up turns to their local government. They request the government to give them access to the DNA data of detainees and convicted offenders. In return, hospitals in the area will get free use of the algorithm, and the government will receive part of profits where the software is sold to third-countries. In this example, detainees and convicted offenders are exploited. They have to give up their privacy and health data (efforts) under the control of the eHealth start-up for the purposes of technology development to the benefit of others (the eHealth, start-up, the local government, as well as the local hospitals and their patients). Exploitation can also occur upon and during the deployment of an algorithmic decision-making process.[878] Gig-economy workers, such as delivery couriers are governed and controlled by

[876] In reference to Macpherson. For a historical and more-in-depth exploration of the concept, see: Young, 48–53.

[877] Crawford and Joler, 'Anatomy of an AI System'.

[878] Even if data have been collected properly, the labelling of data sets can be outsourced upon instruction to workers in low-income countries whose efforts largely go unnoticed. 'AI's New Workforce: The Data-Labelling Industry Spreads Globally | Financial Times', accessed 12 May 2022, https://www.ft.com/content/56dde36c-aa40-11e9-984c-fac8325aaa04. See also on data annotation practices: Miceli, Schuessler, and Yang, 'Between Subjectivity and Imposition'.

automated systems that are designed to optimise the profit of their respective platforms. These systems "reward" efficiency and speed, but lack the working conditions required to mitigate the heightened pressure these systems impose onto workers. The efforts of the delivery couriers are transferred to the benefit of the gig-platform's owners and their end-users who profit from aster delivery services.

Many data-driven systems harvest the data of their end-users and decision-subjects. Under the guise of personalised service provision, data subjects hand over their data, not (only) to their benefit, but to the benefit of those who collect and gather it. Zuboff has defined these practices as surveillance capitalism: "a new (economic) order that claims human experience as free raw material for hidden commercial practices of extraction, prediction, and sales." Surveillance capitalism is an exploitative practice through which the energies of one group (the data of decision subjects) are transferred to those in control over data infrastructures, such as governments and private corporations, to maximise their power over decision-subjects and their data.[879] In this new economic reality, every person who has their data collected as part of data-driven (decision-making) systems can become the subject of exploitation. At the same time, not only the decision-maker stands to profit. By deciding how people will be judged, ranked, classified, or categorised, algorithmic actors establish the conditions that will define who will emerge victorious and who will lose out within a given decision-making context. Decision-making systems can systematically disfavour one group to the advantage of another. Left-handed persons might be denied a loan application in favour of those that are right-handed, or tablet owners might be presented with higher prices so that laptop owners can enjoy better discounts. Based on their information, one person or group will face a particular disadvantage that is transferred as an advantage to others. Without any mechanism of control or insight, such transfers might constitute micro-instances of exploitation.

While data-driven decision-making practices need not be forbidden, decision-makers should carefully reflect upon the transfers of power they are likely to engender. This reflexive exercise could take a moral or categorical imperative as a point of departure: will decision-subjects be seen as an end or solely as a means? In order to address this question, several others might follow: Whose data been used? Who is set to gain the most out of a given decision-making process? Who is set to gain the least? Is there a particular group that is systematically disadvantaged? Do the losses one group incur, transfer as a benefit to others? These questions should not only be asked before or during the design process, but also monitored thereafter. General accountability and

[879] Zuboff, *The Age of Surveillance Capitalism*.

transparency obligations could enable the broader public to question the choices of decision-makers.

POWERLESSNESS – The current risk towards decision-subjects' exploitation can be linked to their lack of power within the digital environment.[880] In the face of the various power asymmetries that characterise the algorithmic environment, decision-subjects (often) lack the authority, status, autonomy and knowledge needed in order to decide upon, or participate in, the definition of the rules, conditions and purposes that will govern their actions or behaviours.[881]

In order to mitigate their exploitation, decision-subjects could be given access to strong regulatory tools that can help them in breaking down these asymmetries. Control mechanisms, such as the aforementioned measures of accountability and transparency can enhance the influence of decision-subjects over their environment. For example, European data protection law tries to empower data subjects. These laws impose additional transparency obligations and give decision-subjects access to several individual rights, including a right to access.[882] This allows decision-subjects to obtain additional insight into the ways their information will be processed and for what purposes, including how their data might be used to the detriment or benefit of others.[883] Moreover, as power is linked to a person's capacity to participate in the governance of their actions and behaviour, inclusive and participatory design approaches can be a valued addition to algorithmic governance strategies forcing decision-makers to take into account the interests and concerns of those they might affect. In the absence of such measures, decision-subjects are bound to remain powerless. Furthermore, general transparency obligations allow a broader public to question the practices of decision-makers. Upon the discovery of an injustice, decision-makers could be called upon to justify their actions and defend the choices they made.[884]

---

[880] Young, *Justice and the Politics of Difference*, 56–58. Power is exercised over the powerless, without their having any capacity to exercise power in return.

[881] See for example: Young, 56 and 219–20.

[882] See for example Articles 12, 13, 14, 15 and 22 GDPR, as well as Section 6.2.2 The Problematic Nature of Generalisations.

[883] Unfortunately, their efficiency in practice would be rendered null if there is no enforcement of their compliance. See for instance: Naudts, Dewitte, and Ausloos, 'Meaningful Transparency through Data Rights: A Multidimensional Analysis'. B.-J. Koops, 'The Trouble with European Data Protection Law', *International Data Privacy Law* 4, no. 4 (1 November 2014): 250–61, https://doi.org/10.1093/idpl/ipu023.

[884] See also: Laurens Naudts, 'Towards Accountability: The Articulation and Formalization of Fairness in Machine Learning', *SSRN Electronic Journal*, 2018, https://doi.org/10.2139/ssrn.3298847.

Algorithmic Marginalisation: Data Exclusion and Over-Exposure

Marginalisation refers to a set of processes through which a person or group is excluded from participating to public life in socially meaningful ways. Due to their exclusion, marginalised groups are rendered invisible and voiceless, unrecognised and isolated, and as a consequence, they face a heightened risk of being deprived from socio-economic benefits, basic rights and liberties.[885] The egalitarian harms produced as a consequence of a decision-maker's failure to adequately recognise the needs and interests of those who are invisible, voiceless and isolated cannot be justified.[886] Within the algorithmic decision-making environment, marginalisation can occur through data exclusion and over-exposure.

DATA EXCLUSION – In marginalisation through exclusion, the injustice lies in a person or group's being barred from participating in valuable, productive, and recognised activities in social structures and relationships of cooperation and organisation. Digital exclusion refers to the (non)deliberate act of excluding an individual or group from digitally participating to algorithmic decision-making systems even though the interests, opinions, data, or information of those excluded could meaningfully inform their design or functioning.[887] For instance, the data set underlying an algorithmic decision-making process might have insufficient information concerning a specific underrepresented group. Due to their exclusion, the system will be less adapted to the particular needs of those underrepresented, and the outcomes produced detrimental to their specific situation.

Consider the Streetbump application, which was designed to passively detect potholes in the city of Boston through accelerometers in drivers' phones.[888] The information passed through by drivers could have been used by the local governments to decide upon the allocation of resources for street reconstruction projects. Crawford rightfully observed how the application had a signal problem: "People in lower income groups [were] less likely to have smartphones, and this [was] particularly true of older residents, where smartphone penetration [could] be as low as 16%. For cities like Boston, this [meant] that smartphone data sets are missing inputs from significant parts of the population — often those who have the fewest resources." Even though older residents could have provided valuable input regarding the construction needs of the city, their lived

---

[885] Young, *Justice and the Politics of Difference*, 53.

[886] See also below: Section 6.1.5 "Cultural Imperialism."

[887] The problem of data exclusion was first introduced by Jonas Lerman, but further modified here. Jonas Lerman, 'Big Data and Its Exclusions', *Stanford Law Review Online* 66, no. 55 (2013), https://doi.org/10.2139/ssrn.2293765.

[888] Crawford, 'The Hidden Biases in Big Data'.

experiences would have been entirely overlooked if the application data would have been relied upon.

Of course, certain persons have an interest in staying invisible. Suppose a public health service develops an online platform for drug addicts to sign up for an alternative medicine programme to help them cope with their dependencies. Drug addicts might not want to share their information through an online portal because they fear this information could be used against them in the future. A person's choice to remain invisible should not be understood as that person's choice to remain voiceless and unrepresented. Their interests still matter, and decision-makers should try to accommodate them to the best of their abilities. The health organisation could contact local civil society and social support groups and set up a system where drug addicts can sign up anonymously to the programme at a physical location. The notion of data exclusion should therefore also extend to instances where decision-makers have failed to consider, or did too little to accommodate, the interests of those who want to remain invisible.[889]

OVER-EXPOSURE – Whereas marginalised people or groups can be excluded, they can also be the subject of over-exposure. Over-exposure occurs in cases where a given person or group in need of a given aid or benefit must endure the arbitrary and invasive authority of those who control access to said benefit. Although a decision-making process can be addressed to overturn or compensate for the (material) disadvantage experienced by a marginalised group (e.g., a welfare program), those labelled as the target of the measure, such as the old, the poor or the disabled, often need to suspend their autonomy and basic rights to privacy and data protection before they can be eligible.[890] Over-exposure should therefore not be confused with inclusive design strategies. The latter have the precise purpose to engage and integrate the views of underrepresented groups in ways that are meaningful and respectful.

In a digital society, over-exposure translates to heightened data surveillance. Ignazio and Klein refer to this problem as the "paradox of exposure", that is, "the double bind that places those who stand to significantly gain from being counted in the most danger from that same counting (or classifying) act."[891] For example, in 1999, the city of Seattle approved a system that would track

[889] Naudts, 'AI En Algoritmische Categorisatie: Gelijkheid En Non-Discriminatie'.

[890] Young observed how "being a dependent in our society implies being legitimately subject to the often arbitrary and invasive authority of social service providers and other public and private administrators, who enforce rules with which the marginal must comply, and otherwise exercise power over the conditions of their lives." Young, *Justice and the Politics of Difference*, 54.

[891] D'Ignazio and Klein, *Data Feminism*, chap. 4, p. 9.

homeless people through a numbering system.[892] Allocated a personal identification number, they entered the "Safe Harbours" program. Their history would be tracked across shelters and programs in order to know them better, serve them better help, allocate them funds, etc. Upon the system's introduction, those affected feared that the system could promulgate discriminatory treatment as well as impose privacy harms. In 2020, the introduction of the SyRI welfare fraud detection system was criticised for its over-surveillance of low-income neighbourhoods.[893]

Over-exposure might also arise over time. For instance, the biased deployment of predictive policing tools has led to so-called runaway feedback loops and/or self-fulfilling prophecies.[894] Certain predictive tools assign a risk score to various districts within a given city. The higher a district's risk score, the higher the perceived risk towards criminality. In areas with a higher risk score, the presence of law enforcement is intensified, and due to heightened police presence more crimes are identified and reported. When those reported cases are also fed into the predictive system, areas with an already high-risk score, see their risk score and the level of surveillance once again increased. In these cases, the wrongful assumption is made that a higher crime identification rate implies a higher rate of criminality. Even though other areas might have the same crime rate, they are never made subject to surveillance. Over time, one area, and all those living within that area, become over-exposed to surveillance.

ALGORITHMIC MARGINALISATION PROCESSES - Algorithmic s**y**stems that are used to mediate the distribution and provision of socially valuable goods, have considerable influence over the way society is structured or organised. In these domains, algorithmic decision-makers have the capacity to exclude (or over-expose) a given person or group on a large-scale and systematic basis from being able to enjoy access to these goods. Whereas these exclusionary harms do not necessarily make it impossible for a person or group to exercise their right to self-development, it might nonetheless hinder their ability doing so. Suppose a popular recruiting company screens incoming applicants on the basis of their hobbies. Their data shows how people with an interest in sports tend to be less loyal to their company. As a consequence, sports fanatics are only given

---

[892] Craig Bicknell, 'Homeless, but ID'd, in Seattle', *Wired*, accessed 23 November 2021, https://www.wired.com/1999/10/homeless-but-idd-in-seattle/.

[893] Vervloesem, 'How Dutch Activists Got an Invasive Fraud Detection Algorithm Banned'. In fact, automated welfare programs often rely upon the collection of data of disadvantaged groups. Eubanks, *Automating Inequality: How High-Tech Tools Profile, Police and Punish the Poor*; Virginia Eubanks, 'The Digital Poorhouse: Permitless Carry and the New Gun-Rights Extremism', *Harper's Magazine*, 1 January 2018, https://harpers.org/archive/2018/01/the-digital-poorhouse/.

[894] See for example: Ensign et al., 'Runaway Feedback Loops in Predictive Policing'.

short-term intern positions, rather than steady job opportunities. Long-term job opportunities might offer people with more certainty in their other choices and life paths.[895] Sports devotees are denied access to a resource that facilitates their meaningful participation in social and public life and could in turn be the starting point of a process towards (further) marginalisation. Given the opacity of algorithms, this process might only manifest itself over, and become visible after, an extended period of time.[896] Considering the cumulative harms that certain exclusionary acts might have on people's socio-relational and socio-economic prospects, the decisional criteria used to differentiate between people, and the design choices going into a system, should be made transparent and subject to (democratic) review and debate before they are put into use in order to avoid that a given person or group (unjustifiably) enters into a process of marginalisation.[897]

Algorithms and Cultural Imperialism

Cultural imperialism universalises a (dominant) group's experience and culture and establishes those experiences and cultures as the norm.[898] A dominant group has the ability to universalise their world views because they have access to the most important "means of interpretation and communication." Through the process of cultural imperialism, the particular perspectives and lived experiences of less privileged groups are obscured. Instead, they are stereotyped and marked out as the "Other."[899] Algorithms constitute a modern means of interpretation and communication. Through their ability to discover and/or apply (new) knowledge, algorithms offer decision-makers renewed ways to interpret and structure society. These new-found interpretations can then be communicated to the outside world as part of algorithms' mediating function. Hence, those in charge of these systems, have the capacity to (unknowingly) shape the world in accordance with their perspective, experiences, and meanings.

WORLD VIEWS AND ALGORITHMS – The decisions that underlie algorithmic systems are value-laden. Classification acts have the power to encode and naturalise a specific ordering of the world.[900] For

[895] See also: Creel and Hellman "Why systematic exclusion matters" in Creel and Hellman, 'The Algorithmic Leviathan', 2021, 37.

[896] See for instance also the negative spiral argument developed in Chapter 1, Section 1.2.2 "Scale and Amplification"

[897] Creel and Hellman, 'The Algorithmic Leviathan', January 2022.

[898] Young, *Justice and the Politics of Difference*, 59.

[899] Young, 59.

[900] The notion classification is used here in a broad sense. It not only refers to the ways in which an algorithmic decision-making process may group together individuals. It also refers to instances where data are classified or labelled in a

example, within computer vision, methods are sought for artificial intelligence to understand the visual world. These algorithms require access to vast amounts of images and videos. These images must be labelled for training purposes, i.e., classified as to what they contain. Whereas certain classification acts might appear trivial and devoid of risk (e.g. does a picture contain a dog or apple), others are anything but.[901] For example, systems that look to constrain, define and classify socially constructed categories, such as gender, race and sexuality, present these categories as detectable and fixed even though they are not.[902] In a similar vein, the definition of highly subjective target variables, such as "the ideal employee", risk gaining an aura of objectivity once they are integrated within a decision-making system.[903] Moreover, some of these classification acts could be seen as inherently demeaning or inferiorising, such as systems designed to classify individuals as either beautiful or ugly.[904] Likewise, algorithmic systems might incorporate stereotyped beliefs concerning a given group. A fraud detection programme might replicate the prejudicial assumption that people living in poverty are more likely to commit fraud.[905] The aforementioned examples illustrate how a subjective (and sometimes unscientific way) of structuring the world can have a detrimental effect for persons or groups that are underrepresented or do not adhere or fit to conventional societal standards.

---

training set. Crawford, *Atlas of AI: Power, Politics, and the Planetary Costs of Artificial Intelligence*, 128 and 139. Crawford and Paglen, 'Excavating AI: The Politics of Images in Machine Learning Training Sets'.

[901] On the role of power in image data sets, see also the work by Milagros Miceli and others: Miceli et al., 'Documenting Computer Vision Datasets'; Milagros Miceli, Julian Posada, and Tianling Yang, 'Studying Up Machine Learning Data: Why Talk About Bias When We Mean Power?', *ArXiv:2109.08131 [Cs]*, 16 September 2021, http://arxiv.org/abs/2109.08131; Milagros Miceli, 'AI's Symbolic Power: Classification in the Age of Automation', 2019; Milagros Miceli and Julian Posada, 'A Question of Power: How Task Instructions Shape Training Data', https://sites.google.com/sheffield.ac.uk/bhcc-2020/program?authuser=0; Miceli, Schuessler, and Yang, 'Between Subjectivity and Imposition'.

[902] Crawford, *Atlas of AI: Power, Politics, and the Planetary Costs of Artificial Intelligence*, 144.

[903] Crawford, chaps 4, Classification.

[904] Even in the case of opaque decision-making systems, where decision-subjects do not know how or on what grounds they are classified, a decision-maker should still realise that such a system only serves to demean a given group. See for example: Arwa Mahdawi, 'This AI-Powered App Will Tell You If You're Beautiful – and Reinforce Biases, Too', *The Guardian*, 6 March 2021, sec. Opinion, https://www.theguardian.com/commentisfree/2021/mar/06/ai-powered-app-tell-you-beautiful-reinforce-biases. Sam Levin, 'A Beauty Contest Was Judged by AI and the Robots Didn't like Dark Skin', *The Guardian*, 8 September 2016, sec. Technology, https://www.theguardian.com/technology/2016/sep/08/artificial-intelligence-beauty-contest-doesnt-like-black-people.

[905] 'Profiling the Poor in the Dutch Welfare State - NYU School of Law – CHRGJ', accessed 3 August 2022, https://chrgj.org/2019/11/01/profiling-the-poor-in-the-dutch-welfare-state/; Huisman, 'SyRI, het fraudesysteem van de overheid, faalt'; van Bekkum and Borgesius, 'Digital Welfare Fraud Detection and the Dutch SyRI Judgment'.

Those not discretely captured by a given classification act are rendered invisible and designated as an Other. Once a system is deployed, those value-laden choices gain traction, and their impact can be lasting and profound. For instance, scanning algorithms at airports have misidentified transgender and gender nonconforming persons and wrongfully categorised them as a security threat. These errors occur because these systems have been built with a specific, non-inclusive, non-universal definition of the human body or gender in mind.[906] Such systems obscure the lived physical and psychological experiences of those involved. Moreover, when the security alarm goes off, this might trigger an invasive security check. In this way algorithms contribute to the violation of a person or group's physical integrity.[907]

Algorithmic decision-makers always run the risk of *projecting* their own experience as representative of reality, thereby overlooking the lived experiences of others.[908] They should therefore be encouraged to reflect upon, and make explicit, the assumptions that underlie their decision-making processes, and consider more profoundly the impact their choices have, whether material or non-material, physical or psychological.[909] Reflexive self-monitoring obligations can generate awareness among those in control over the means of interpretation and communication. At the same time, the outcomes of these exercise can be made subject to review.[910] Yet, the overrepresentation of a decision-maker's subjective experiences might only be thwarted when they are confronted with a plurality of views. Diversity, inclusion, participation and careful reflection during the design, deployment and monitoring of decision-making systems are necessary steps to ensure the democratisation of the modern means of interpretation and communication, which too often lie in the hands of a few select actors.[911]

---

[906] Medina, 'TSA's Body Scanners Are Gender Binary. Humans Are Not.'

[907] Similar misidentification issues have also been experienced by people with disabilities, people who wear religious headgear and women of colour. Medina.; Frank, 'TSA Agents Say They're Not Discriminating Against Black Women, But Their Body Scanners Might Be'.

[908] Young, *Justice and the Politics of Difference*, 59.

[909] See among others, the many sources in: Miceli et al., 'Documenting Computer Vision Datasets'. Miceli, 'AI's Symbolic Power: Classification in the Age of Automation'; Miceli and Posada, 'A Question of Power: How Task Instructions Shape Training Data'; Miceli, Schuessler, and Yang, 'Between Subjectivity and Imposition'.

[910] Miceli et al., 'Documenting Computer Vision Datasets'.

[911] See Young Young, *Justice and the Politics of Difference*, 59.

Algorithms and Violence

Oppressed social groups may suffer the oppression of systemic violence.[912] While individual acts of violence, whether they be physical or psychological, constitute a moral wrong, violence as oppression is a social phenomenon: it is the totality of social conditions that render individual acts of violence possible, acceptable, and expected. Violence is a systemic, social practice. It is aimed directly towards, or experienced by, members of a group because of their group membership and those who are oppressed have a clear understanding that they are liable to face the threat of violence.[913] At the same time, the violence they experience, and fear is tolerated by society at large and its occurrence therefore gains a sort of legitimacy.[914]

Algorithmic decision-making processes can reflect or replicate the systematic threat towards violence that social groups experience and cause direct physical and psychological harm to their group members.[915] Insomuch as people of colour fear the threat of unjust police violence, they may also fear the deployment of predictive policing algorithms. Despite the latter showing a higher error rate towards people of colour, these systems nonetheless remain in use. These data-driven systems expose people of colour to a heightened risk of unduly police interference.[916] Refugees may experience intense stress upon the discovery that they will face invasive emotion detection technologies during border control investigations.[917] As above-mentioned, transgender

---

[912] Young, 60.

[913] Young, 62.

[914] Young, 61–62.

[915] In this context, Article 5 (b) of the Proposed AI Act recognises the physical and harm algorithmic systems could cause, and wishes to prohibit "the placing on the market, putting into service or use of an AI system that exploits any of the vulnerabilities of a specific group of persons due to their age, physical or mental disability, in order to materially distort the behaviour of a person pertaining to that group in a manner that causes or is likely to cause that person or another person physical or psychological harm."

[916] See for example: Kashmir Hill, 'Another Arrest, and Jail Time, Due to a Bad Facial Recognition Match', *The New York Times*, 29 December 2020, sec. Technology, https://www.nytimes.com/2020/12/29/technology/facial-recognition-misidentify-jail.html. Christina Pazzanese Harvard Staff Writer, 'How Unjust Police Killings Damage the Mental Health of Black Americans', *Harvard Gazette* (blog), 13 May 2021, https://news.harvard.edu/gazette/story/2021/05/how-unjust-police-killings-damage-the-mental-health-of-black-americans/; Jocelyn R. Smith Lee and Michael A. Robinson, '"That's My Number One Fear in Life. It's the Police": Examining Young Black Men's Exposures to Trauma and Loss Resulting From Police Violence and Police Killings:', *Journal of Black Psychology*, 30 July 2019, https://doi.org/10.1177/0095798419865152.

[917] See for example: Javier Sánchez-Monedero and Lina Dencik, 'The Politics of Deceptive Borders: "Biomarkers of Deceit" and the Case of IBorderCtrl', *Information, Communication & Society* 25, no. 3 (17 February 2022): 413–30, https://doi.org/10.1080/1369118X.2020.1792530.

and gender nonconforming persons may experience a sense of dread upon seeing airport body scanners.[918] Facebook's algorithms have been accused of making hate worse.[919] In all the aforementioned cases, the sole deployment of algorithmic systems could already be a sufficient cause for psychological harm. Yet, the decisions that these systems produce could lead to further physical and/or psychological harm.

While non-oppressed individuals or members of non-representational, algorithmic aggregates may also experience physical or psychological harms as a result of flawed (or unfavourable) decision-making systems, they do not face the oppression of violence. Due to their lack of social cohesion and dynamic nature, the harmful acts that algorithmic aggregates and their group members may endure, are unlikely to gain a systemic, socially tolerated, accepted, and expected character. Hence, violence in particular is a socio-relational harm that can only be experienced by historically disadvantaged or social groups.

6.1.6 Algorithmic Domination[920]

Domination constrains people in their right to self-determination and autonomy.[921] Social justice requires that people are free from domination in their private relationships (including their relationships with private corporations), whereas political legitimacy requires governments to refrain from public domination.[922] Algorithmic domination more generally refers to the capacity of an algorithmic decision-maker, both public and private, to interfere with a person (or group)'s

---

[918] See for example: Keyes, 'The Misgendering Machines'; Medina, 'TSA's Body Scanners Are Gender Binary. Humans Are Not.'

[919] 'Frances Haugen Says Facebook Is "Making Hate Worse"', *BBC News*, 25 October 2021, sec. Technology, https://www.bbc.com/news/technology-59038506.

[920] This section draws upon Pettit's definition and work on non-domination in order to define algorithmic domination. The following works in particular have been important: Pettit, *Republicanism*; Pettit, *On The People's Terms: A Republican Theory and Model of Democracy*.

[921] Pettit himself has developed a rich or substantive understanding of autonomy as "orthonomy": "the capacity to form one's desires according to the right sort of principles rather than the wrong." Philip Pettit and Michael Smith, 'Backgrounding Desire', *The Philosophical Review* 99, no. 4 (October 1990): 565, https://doi.org/10.2307/2185617; Pettit, *Republicanism*, 81; Alexa Nossek and Julia Belz, 'Notes on Pettit's Concept of Orthonomy', in *Philip Pettit: Five Themes from His Work*, ed. Simon Derpmann and David P. Schweikard, Münster Lectures in Philosophy (Cham: Springer International Publishing, 2016), 61–69, https://doi.org/10.1007/978-3-319-26103-4_6.

[922] Pettit describes his theory as expressively egalitarian: "A republican theory of justice would seek the expressive equalization of freedom as non-domination: the promotion of freedom as nondomination on the basis of an equal concern for each citizen." Pettit, *On The People's Terms: A Republican Theory and Model of Democracy*, 3. Pettit, 298.

(valuable) choice options in a manner that is arbitrary, unchecked or uncontrolled.[923] The evaluation of algorithmic decision-making systems can therefore be led by the question whether they are likely to increase (or decrease) a person or group's domination. At the same time, regulatory initiatives could explore how the algorithmic environment is best structured in order to bolster freedom as non-domination: what resources do decision-subjects need access to in order exercise free choice and how can they be protected against the arbitrary interference of others in their exercise of it?[924]

In the following sub-Sections, the notion of algorithmic domination will be further unpacked. The ideal of freedom of non-domination promotes people's enjoyment of free or undominated choice.[925] When a person has a choice between a set of options, their freedom can be reduced through the hindrance of any of those options.[926] In order to understand how dominance can arise in the relationship between decision-maker/decision-subject, it important to understand what it means for people to have a choice and how their choices can be interfered with in a manner that is uncontrolled or unchecked. In order to map how decision-makers can increase a decision-subjects' vulnerability vis-à-vis others, it will also be important to ascertain the resources people require in order to exercise choice.

---

[923] See for instance, Pettit's definition at: Pettit, *On The People's Terms: A Republican Theory and Model of Democracy*, 50.

[924] Pettit, 69.

[925] Pettit, 82.

[926] Pettit states the following: "I propose to understand freedom in such a way that in a choice between given options, say X, Y and Z, or just X and Y, your freedom can be reduced by the hindrance of any one of those options, regardless of which you happen to prefer. Such a hindrance will make one or more of the options unavailable, undermining your objective or cognitive resources." The options are available insofar as they are objectively true (a person *can* do X, Y or Z and nothing else instead), and this truth is cognitively registered (a person *understands* and *realises* they *can* choose X, Y, or Z and nothing else instead). Pettit, 26–33. Freedom as non-domination does not distinguish between preferred and unpreferred options. This means that a person will also be hindered in the exercise of their choice when the hindrance affects an option they would not have chosen. Pettit follows Berlin in arguing that any hindrance in choosing between a set of options should be considered a loss of freedom. This view is contrasted to Hobbes, who argued that this can only be said of hindrances to preferred options. Yet, focusing on "preferred" options only can give rise to some counter-intuitive circumstances. For instance, under a Hobbesian account, freedom could be accompanied by force, if the choice forced upon me corresponds to my actual preference, there is no problem. In a similar vein, people would become free by adapting their preferences, even if they would adapt themselves as a result of constraints (See also Nussbaum's observations on adaptive preferences in Chapter 7). Pettit, ch 1.

What Choice?

Within the algorithmic environment, people, on many occasions, are presented with a range of options over which they can exercise a choice. Movie platforms give people options on what they can watch on their television sets. Social media platforms suggest who they can follow and interact with or what news articles they can read. Public unemployment agencies offer the unemployed a variety of job applications to consider. Would all these instances of choice necessitate freedom from non-domination? Though it is for regulators to decide (following public debate) what areas of choice should remain undominated, some general points of guidance can be formulated.

SAFEGUARDING FUNDAMENTAL CHOICES AND FREEDOMS – Under a first approach, the decision could be made to maximise people's undominated status over a range of choices related to their fundamental interests, basic liberties or other justice-relevant (non)material goods or services.[927] Once these domains of choice have been identified, regulators should limit the capacity of algorithmic decision-makers' to dominate decision-subjects in their enjoyment thereof. On the one hand, in order to ensure people are free in their exercise of a choice, regulators would need to resource or facilitate choice. On the other hand, regulators ought to protect people in their exercise of choice. Suppose access to news information has been identified as a justice-relevant good. Nowadays news curation has become increasingly mediated by a select few social media platforms who control how news is presented and consumed. If people ought to enjoy non-domination over their news media consumption, regulators should give decision-subjects the tools they need to regain control over their media feeds. For example, they could mandate news platforms to make clear and accessible the main parameters they use within their recommender systems, as well as provide options to decision-subjects to modify their influence.[928] At the same time, the law could prohibit media platforms to alter people's news options unilaterally or covertly.

Certain algorithmic structures have such a profound impact on the way people exercise their fundamental life's choices and rights that their dominating character is almost unquestionable. Mass-surveillance programs have a chilling effect on the way people enjoy and exercise their

---

[927] Pettit for example believes that in order to enjoy freedom of non-domination over a range of choices comprised by the basic liberties (e.g., the freedom to think what you like or the freedom to change occupation and employment). People should have an undominated status, guaranteed through resources and protection, such as laws that entrench these liberties. Pettit, *On The People's Terms: A Republican Theory and Model of Democracy*, chaps 2, Social Justice.

[928] See for example: Article 29, § 1 Proposed DSA.

fundamental rights to privacy, data protection, freedom of speech, etc.[929] Social sorting programs act in a similar directly dominating manner.[930] In both cases, algorithmic techniques are layered as a membrane onto society in order to track and monitor our daily activities, and if need be, interfere in those actions in highly dynamic, adaptive and/or autonomous ways.[931] It is therefore not unsurprising that the call towards their abolition has been the greatest.[932]

The ability to interfere into people's fundamental choices need not be linked to massive algorithmic structures or tied to areas where algorithms guide critical decisions. Even in the case a decision-maker exerts control over a set of (seemingly) trivial options, their influence and power could still be considerable enough for them to render certain decision-subjects, through among others acts of oppression, more vulnerable towards domination. Suppose a specific digital movie platform has a monopoly on the way people consume entertainment at home. The platform makes use of a recommender system that perpetuate gender stereotypes when selecting movies to end-users. The platform systematically recommends women cooking shows and men hard science documentaries. By reinforcing a traditional distribution of gender roles, the platform helps maintain the dominated and oppressed status of women, which in turn can restrict their freedoms to areas of care and family life.

MAXIMISING SELF-DETERMINATION – As a policy goal, freedom as non-domination could also be constructed as an objective to maximise decision-subjects' right to self-determination in their direct relationship with algorithmic decision-makers. Under this second approach, regulators

---

[929] Gräf, 'When Automated Profiling Threatens Our Freedom', 446; Muldoon and Raekstad, 'Algorithmic Domination in the Gig Economy'. J. Matthew Hoye and Jeffrey Monaghan, 'Surveillance, Freedom and the Republic', *European Journal of Political Theory* 17, no. 3 (July 2018): 343–63, https://doi.org/10.1177/1474885115608783.

[930] See for instance: 'China "Social Credit": Beijing Sets up Huge System - BBC News', accessed 14 September 2019, https://www.bbc.com/news/world-asia-china-34592186; Axel Dorloff and Daniel Satra, 'China: Auf dem Weg zur totalen Überwachung', *tagesschau.de*, 24 March 2019, https://www.tagesschau.de/ausland/ueberwachung-china-101.html.

[931] Of course, these systems are always deployed *by* someone or some agency. An algorithmic system may act autonomously through learning, but the choice of maintaining a system like this has been made by a person or government.

[932] Article 5 (c) of the Proposed AI Act intends to prohibit social sorting systems: "the placing on the market, putting into service or use of AI systems by public authorities or on their behalf for the evaluation or classification of the trustworthiness of natural persons over a certain period of time based on their social behaviour or known or predicted personal or personality characteristics, with the social score leading to either or both of the following: i) detrimental or unfavourable treatment of certain natural persons or whole groups thereof in social contexts which are unrelated to the contexts in which the data was originally generated or collected; ii) detrimental or unfavourable treatment of certain natural persons or whole groups thereof that is unjustified or disproportionate to their social behaviour or its gravity.

would resource and protect decision-subjects against arbitrary interference when decision-makers have direct control over the former's options and choices. Gräf for example has argued that algorithmic structures that limit, replace or create the options of decision-subjects (based upon their data, behaviour or actions), such as video portals or search engines, all have the potential to become instances of unwarranted domination if end-users have no deliberative degree of control over the way their scope of action is adapted.[933]

Following this approach, a broad range of algorithmic decision-making systems would be captured by the demands of non-domination. Of course, these systems need not be problematic *per se*. Much will depend upon the reasonable and rightful expectations decision-subjects have when they enter into algorithmic decision-making relationships and the level of protection and control they are offered within them. As above-mentioned, decision-subjects could be given clear and transparent information on how their options can be curated, personalised, ranked, streamlined, or altered, or given insight and control over how their information can be used and for what purposes.[934] Information requirements and obligations could thus enable decision-subjects to deliberately enter into an algorithmic relationship on equal footing. Knowing their choices will be interfered with, these interferences are no longer arbitrary, but willed.[935] When a decision-maker has the unilateral and ungoverned capacity to (covertly) change how decision-subjects exercise their choice, or decision-subjects have been misinformed about the level of control the decision-maker has, these could constitute instances of domination. The latter illustrates the supportive function procedural fairness safeguards hold for maintaining egalitarian social relationships. In fact, consumer protection and data protection laws often serve this exact purpose: they look to govern information flows in such a way that decision-subjects can calibrate their expectations in relation to their digital surroundings in order to break down detrimental asymmetries in power.[936]

---

[933] Gräf, 'When Automated Profiling Threatens Our Freedom', 450.

[934] For instance, a decision-subject might want to know whether their information can be used to oppress or exclude other groups from enjoying a specific benefit or interest.

[935] The ability for decision-subjects to have access to alternatives can be important as well. If a given actor has a monopoly on a given service or good, decision-subjects will have no choice but to enter into a relationship with them. These positions of competitive power also increase the risk of domination.

[936] In this context one could argue that the fundamental interest people have to have their right to self-determination maximised would become the justice-relevant good or basic liberty. Similarly, one could argue that maximising the right to self-determination protects decision-subjects in how they want to exercise choice over their enjoyment of among others the fundamental rights to privacy and data protection.

NO CHOICE? – Algorithmic decision-makers can govern a person's freedom to choose over a range of options incidentally, rather than directly. When a financial institution makes use of a credit scoring algorithm to grant or deny credit applications, this decision does not constitute an interference with people's options. As Gräf points out, whether the decision-subject will obtain a loan is not their choice to make.[937] Likewise, algorithmically guided public decisions, such as a welfare fraud detection programme or recidivism risk estimation, do not offer decision-subjects a range of options to choose from. Still, even in these situations, the ideal of non-domination holds importance. First, decision-subjects might still carry reasonable expectations as to how these systems will function. Suppose a recruiting company screens the social media profiles of applicants through automated means. Non-discrimination laws prohibit discrimination on the basis of race and ethnicity in the employment context. If the system appears biased on the grounds of ethnicity, one could argue that the decision arbitrarily interferes into a person's life choices. A person might not have a choice in their getting a job position, but the law does create a rightful expectation that as part of the recruiting process race and ethnicity are not decisive. Second, in certain cases, the decisions algorithms mediate, like credit and job applications, affect, what Creel and Hellman refer to as people's realms of opportunity.[938] When a person or group are denied a loan or job position, it diminishes their access, and hence freedom, over a range of (future) choices. In these contexts, regulators should carefully consider the positions of power decision-makers occupy, and the conditions under which these decisions are made. The use of algorithms to guide these decisions need not be prohibited but because they have a greater capacity to render persons or groups more vulnerable, the design choices and decision criteria relied upon to differentiate between people, might benefit from a (democratic) review prior to their use. [939]

---

[937] See also: Gräf, 'When Automated Profiling Threatens Our Freedom', 447.

[938] Kathleen Creel and Deborah Hellman, 'The Algorithmic Leviathan: Arbitrariness, Fairness, and Opportunity in Algorithmic Decision Making Systems', in *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency*, FAccT '21 (New York, NY, USA: Association for Computing Machinery, 2021), 816, https://doi.org/10.1145/3442188.3445942. See also Section 6.1.5 Algorithmic Marginalisation.

[939] Decision-makers ought to be encouraged to actively reflect upon the impact they have on people's choices and vulnerability status, and to monitor the practical consequences their technologies might have on people's undominated status. From a regulatory perspective, one should nonetheless remain pragmatic, and acknowledge that not every actor or action presents the same type of risk towards domination. Such a gradual approach could be likened to a risk-based approach found in certain regulations. For example, the GDPR and proposed AI Act heighten a given actor's accountability and responsibility depending upon the risk their activities present for the fundamental rights and freedoms of those involved. Whereas the GDPR states that every data controller must respect data quality principles, in the case their activities pose a heightened risk towards the fundamental rights of people, they are also obliged to perform a data protection impact assessment. See e.g., Article 35 GDPR.

Interference in (the resources) of Choice

People can only exercise free choice over a range of options when they have the resources to do so. Pettit identifies three sets of resources. Personal resources relate to the mental and bodily capacities needed to make a choice. Natural resources relate to the environmental conditions that make it actually possible to act upon a given option. Finally, social resources relate to societal conditions and conventions that allow a given choice to be made. As aforementioned, these resources can include among others the access, and enjoyment of, fundamental choices, basic liberties or other justice-relevant goods, legal protection, wealth and income, the certainty of having a steady job position, the ability to obtain a loan, etc.

Algorithms are both a social and natural resource for choice. They are social as they take instructions and values from their surroundings and upon deployment. In addition, they have the capacity to co-shape new social meanings. Algorithms are a natural resource in the sense that they are technological and environmental artefacts that decision-subjects need to interact with in their navigation of everyday life. Algorithms are a resource because they can either enhance or limit people in the exercise of choice, including their access to other resources for choice. Drawing from Pettit's distinction between vitiating and invading hindrances, algorithms can interfere with people's resources for choice in two specific ways.

First, an algorithmically guided decision-making process can affect people's resources for choice in a general way **(vitiating hindrances)**. Hiring and credit scoring algorithms for instance change people's socio-economic prospects, and as a consequence, the range of choices they will have access to. Algorithms can thus expand or hinder people in the enjoyment of resources that define their "range of effective opportunities."[940]

Second, certain algorithmic decision-making systems are built in such a way that they only allow decision-subjects to exercise choice over a range of options within the limits or conditions decided upon by the decision-maker **(invading hindrances)**. Platforms can make use of behavioural analytics to keep decision-subjects active and engaged, which empowers decision-makers to maximise their business interests. In this case, they can lock-in decision-subjects to increase the profits they obtain from targeted advertisements. The decision-maker usurps control over the decision-subjects' choices. They are unable to exercise choice in accordance with their preferences, or they can do so only within the limits dictated by the decision-maker. Pettit

[940] Pettit, *On The People's Terms: A Republican Theory and Model of Democracy*, 45.

describes the difference thusly: "To suffer the vitiation of choice is to be denied a precondition for enjoying freedom of choice: to lack required resources. To suffer invasion is to be denied the very condition by which freedom is identified: to be thwarted in making the choice according to your will."[941] Only invading hindrances can constitute (direct) domination on the part of the decision-maker.[942]

Active and Passive Domination

Presuming a decision-maker governs a protected range of choices, the following paragraphs will explore how, in the vertical relationship decision-maker/decision-subject, the former can dominate the latter. They moreover concern situations where decision-subjects have no democratic or deliberative control over the algorithms' operations and functionings, and the decision-maker's actions are not directed, dictated, or outsourced on the request of decision-subjects in a deliberative way.

ACTIVE DOMINATION - A decision-maker can dominate decision-subjects through direct action. A first type of active interference concerns the possibility for the algorithmic decision-maker to remove, replace, or attach a penalty, to a person's options.[943] For example, decision-makers can deploy learning techniques to (dynamically) alter the options presented to decision-subjects

---

[941] Pettit, 44. The distinction between invading and vitiating hindrances, corresponds to the distinction between vertical and horizontal domination, respectively.

[942] Pettit, 37–38. Invading hindrances are inherently inimical to free choice. Vitiating hindrances are inimical in an incidental manner. Invaders are triggered by a person's (prospective) attempt to exercise choice. Vitiators may materialise independently of a person's attempt to exercise choice. For more information on these two notions, see also: Pettit, chaps 2, 37–47.

[943] The addition of an option is not problematic as long as decision-subjects can maintain a degree of control over the way they exercise choice over the added options. Similarly, adding a reward to a given option should not be considered as replacing an option as long as it remains possible for the decision-subject to refuse the award. This might not be the case however where a person is flooded with additional options. The latter might constitute a form of manipulation. As Pettit observes: "manipulation may involve […] snowing you with so much information that you are putty in my hands." This is for instance a known problem in consumer protection and data protection scholarship: dense privacy policies discourage consumers and data subjects to act against traders and data controllers' unfair business practices. Yet, within algorithmic decision-making systems, the decision-subject might be presented with as many options that they are manipulated into choosing an option that is most-preferred by the decision-maker. For example, a search engine might return so many options that the decision-subject chooses those that are presented to them first. This search result might have however been produced by the preferred partner of the search engine, and as such manipulates the decision-subject to act in accordance with the will of the search engine.
Pettit, *On The People's Terms: A Republican Theory and Model of Democracy*, 51–52.

based on their actions, behaviours or inferred preferences and habits. Suppose a social media platform has an interest to support one of two presidential candidates. Candidate A wants to deregulate the tech industry, whereas candidate B is in favour of more strict rules. Through the use of inferential analytics, the platform tries to detect which end-users would be in favour of candidate B. When those decision-subjects navigate the platform, positive news bulletins about B will be replaced by slanderous ones. Not only does the platform directly control the news options decision-subjects are presented, they also directly interfere with people's freedom to think, express and associate with certain political opinions.

A second type of direct hindrance concerns the algorithmic decision-maker's capacity to misrepresent the options a person has available through strategies of deception and manipulation.[944] Inferential techniques may not only be used to (dynamically) alter a person's options. The knowledge they attain can often be exploited to deceive or manipulate decision-subjects into choosing or performing a given action.[945] By only presenting slanderous news bulletins about presidential candidate B, decision-subjects were manipulated into choosing the candidate the social media platform preferred. An online storefront could combine preference detection with emotion analytics. If they detect that a decision-subject overestimates the utility of a product they prefer, they could steer them into buying the product at supra-competitive prices.[946] Many data-driven techniques could be used to achieve such a "nudging" effect. Due to their highly personalised, dynamic, adaptive, and pervasive nature, combined with the complexity and opacity of the digital environment in which they operate, such effects may remain hidden however, which in turn renders the use of these steering tactics more invidious.[947]

---

[944] Deception involves the strategy where decision-subjects are deceived into thinking the options available are different from what they are in reality, or that the consequences attached to choosing a given option are different. Manipulation also changes how choices are made. Manipulation denies someone the possibility to choose between options "on the basis of a proper understanding of the options on offer." Pressure is applied onto the decision-subject so they cannot reflect clearly about the options they have been given. Pettit, 54–56.

[945] For additional examples on digital manipulation, see also: Philipp Hacker, 'Manipulation by Algorithms. Exploring the Triangle of Unfair Commercial Practice, Data Protection, and Privacy Law', *European Law Journal* n/a, no. n/a, accessed 10 May 2022, https://doi.org/10.1111/eulj.12389. Damian Clifford, 'Citizen-Consumers in a Personalised Galaxy: Emotion-Influenced Decision Making, a True Path to the Dark Side?', in *Future Law: Emerging Technology, Regulation and Ethics*, ed. Edina Harbinja, Lilian Edwards, and Schafer Burkhard (Edinburgh University Press, 2020), 79–130, https://www.cambridge.org/core/books/future-law/citizenconsumers-in-a-personalised-galaxy-emotioninfluenced-decision-making-a-true-path-to-the-dark-side/8EC23988B488888E1F8DC5FD15FF17D3.

[946] See also: Hacker, 'Manipulation by Algorithms. Exploring the Triangle of Unfair Commercial Practice, Data Protection, and Privacy Law'.

[947] This is what Yeung referred to as "hypernudges." Karen Yeung, '"Hypernudge": Big Data as a Mode of Regulation by Design', *Information, Communication & Society* 20, no. 1 (2 January 2017): 118–36,

Admittedly, algorithmic techniques, such as inferential analytics and recommender systems, might also be used to maximise decision-subjects' freedom of choice in their navigation of the online world.[948] They can have a streamlining function and point people towards those options that are of most interest to them. The distinction between enhanced choice and dominated choice is therefore not always easily made.[949] When algorithms are used to (predominantly) satisfy the interests decision-makers, they still enable decision-makers to claim a share in decision-subjects' control over choice. Though beneficial, algorithms would nonetheless establish an inegalitarian social relationship if decision-subjects are left without deliberative insight or control over them.

PASSIVE DOMINATION - Even in the case decision-makers do not actively interfere with the options of decision-subjects, the former could still be in a position of domination towards the latter.[950] There are two broad ways through which passive domination can occur: invigilation and intimidation.[951] Invigilation refers to the situation where the algorithmic decision-maker stands guard over their decision-subjects: as soon as they desire to, they can readily interfere with their choices.[952] Due to their immediate access to the infrastructures, decisions and (underlying) knowledge, most decision-makers hold this power. They can unilaterally change the code or underlying logic of an algorithm, changing the options the decision-subject would typically have access to. They might not exercise this power, but if they were to change their disposition, they easily could. When decision-subjects become aware of this capacity, they could experience this as intimidating. Decision-subjects may also be intimidated when they believe the decision-maker has the uncontrolled capacity to interfere in their options, even though the decision-maker might not have this power in practice.[953]

---

https://doi.org/10.1080/1369118X.2016.1186713; Danaher, 'Freedom in an Age of Algocracy', 265. Interestingly, Pettit too considered the permissible use of nudges. These practices could be tolerated under controlled conditions and if their impact on choice remains limited. Pettit gives the example of organ donation. Persons are more likely to agree to organ donation upon their license registration when the question is presented to them in the form of an opt-out as opposed to an opt-in. He concludes that "The opt-out system probably evokes more positive choices because it communicates the presumptively correct and non-deceptive message that there are no horrific consequences attached to the choice; if there were, the opt-in system would scarcely have survived. The nudging involved does not replace any of the options on offer with a distorted counterpart and it does not have the potential to affect a person's freedom in making the choice between signing up and not signing." Pettit, *On The People's Terms: A Republican Theory and Model of Democracy*, 56, at footnote 32.

[948] See for instance, Danaher on "choice filtration" and "cognitive slack": Danaher, 'Freedom in an Age of Algocracy'.

[949] Gräf, 'When Automated Profiling Threatens Our Freedom', 448.

[950] Pettit, *On The People's Terms: A Republican Theory and Model of Democracy*, 59–60.

[951] Pettit, 61.

[952] Pettit, 61.

[953] Pettit, 61–62.

Adaptive learning systems stand guard over decision-subjects. Options might be displayed in a specific fashion, but as soon as a point of interest is discovered by data-driven analytics, such as a decision-subject's preferences or habits, a corresponding action can be triggered. A social media platform for example can give third parties the possibility to run targeted advertising campaigns. While decision-subjects may feel they can navigate the platform freely and under their control, as soon as a third party discovers a decision-subject belongs to a desired demographic, their regular options are changed into sponsored ones.

The deployment of algorithmically guided decision-making systems can also be experienced as intimidating. Oppressed social groups for example might fear the use of inferential analytics because they allow others to uncover sensitive information they did not (yet) want to share. For example, a person may want to keep their sexual orientation hidden from family members. They might worry that once that information has been inferred, they can be exposed through targeted advertisements displayed on the shared family computer. Airport body scanners can be intimidating for gender-nonconforming persons knowing these systems might misidentify them. Both examples illustrate how intimidating algorithms can interfere with a person's freedom to choose. In the first case, intimidation removed a person's option to freely navigate the internet, in the second the option to travel abroad.

The idea that decision-makers can shape decision-subjects options upon their sole discretion, or make use of novel tools that grant expanded insight into a person's thoughts, actions, or emotions, might be more generally experienced as intimidating however, and could in itself affect how decision-subjects exercise their choice both off-and online.[954] People may alter the way they present themselves online, because they believe their current or future employer makes use of social media analytics to monitor their behaviour. If this fear does indeed exist among the broader public, this would indicate their vulnerability vis-à-vis algorithmic decision-makers and it would signal a regulatory environment incapable of guarding people against socio-egalitarian status-harms.

[954] See also: Moritz Büchi et al., 'The Chilling Effects of Algorithmic Profiling: Mapping the Issues', *Computer Law & Security Review* 36 (April 2020): 105367, https://doi.org/10.1016/j.clsr.2019.105367.

The question can be raised whether algorithmic structures, rather than isolated decision-making instances, can exert dominating power.[955] Mass surveillance and social scoring algorithms for example stand guard and intimidate through a variety of interconnected processes.[956] At the same time, people' choices and decisions are steered and interfered with in many little and often invisible and opaque ways, through the (isolated) actions of many algorithmic decision-makers. While these interactions and their effects may often go unnoticed, they can nonetheless become tangible once we try to deviate or escape from their grasp. Danaher explains how the algorithmic

---

[955] Pettit argues a distinction may need to be made between direct or personal types of domination on the one hand, and indirect or structural forms of domination on the other hand, which arise (among others) through socio-relational and socio-economic inequalities: "These modes of organization may vitiate, but not invade, choice […], but they can indirectly facilitate the worst forms of invasion and domination in a society." Pettit, *On The People's Terms: A Republican Theory and Model of Democracy*, 63. In reference to Hayward 2011. This is one point where social egalitarian authors believe that Pettit's notion of non-domination could be further strengthened. A first core critique takes aim Pettit's description of domination as an interpersonal and agential phenomenon. Under this conception, domination takes place between an identifiable agent (or agency) who exerts uncontrolled power over another agent. A second core critique concerns the interpretation of the concept interference as needing to be intentional. A narrow understanding of non-domination may therefore fail to capture more structural forms of domination, oppression, and social marginalisation, which often come about through the unintentional or unconscious actions of many people. If the capacity to dominate is also seen as a structural phenomenon which permeates society, the notion could capture a greater variety of socio-relational harms, as well as a greater variety of instances where such harms could occur. Modifying domination this way could also benefit the evaluative function non-domination holds within the algorithmic environment. First, where algorithms replicate the structural and systemic forms of discrimination and disadvantage, the notion of non-domination could still be used to designate these negative consequences as an unjustified, socio-relational egalitarian harm. Second, non-domination could more easily capture the harmful effects that flow from the algorithmic environment as a structure, such as algorithmic micro-domination or indirect horizontal domination, and explain why these phenomena threaten social equality too. For a more thorough analysis of the social egalitarian critique, see also: Garrau and Laborde, 'Relational Equality, Non-Domination, and Vulnerability'; Schuppert, 'Non-Domination, Non-Alienation and Social Equality'; Sharon R Krause, 'Beyond Non-Domination: Agency, Inequality and the Meaning of Freedom', *Philosophy & Social Criticism* 39, no. 2 (1 February 2013): 187–208, https://doi.org/10.1177/0191453712470360; Michael J. Thompson, 'Reconstructing Republican Freedom: A Critique of the Neo-Republican Concept of Freedom as Non-Domination', *Philosophy & Social Criticism* 39, no. 3 (March 2013): 277–98, https://doi.org/10.1177/0191453712473081; Michael J Thompson, 'The Two Faces of Domination in Republican Political Theory', *European Journal of Political Theory* 17, no. 1 (January 2018): 44–64, https://doi.org/10.1177/1474885115580352.

[956] As Danaher observes: "If your behaviour is being monitored and mined for patterns and predictions, then it is possible that some of that behaviour might trigger interference from the system itself (thanks to automation) or from some human controller of the system, particularly if it falls outside the normal or permissible range of the system's expectations." Danaher, 'Freedom in an Age of Algocracy', 266. Gräf, 'When Automated Profiling Threatens Our Freedom', 446; Danaher, 'Freedom in an Age of Algocracy', 266. Hoye and Monaghan, 'Surveillance, Freedom and the Republic'.

environment might therefore be comprised of various instances of micro-domination.[957] Because all of our actions are captured, mediated or governed in the digital environment, the attempts people make to escape their grasp, can be considered as suspicious, criminal or otherwise punishable.[958] In so doing, algorithms "create a space of permissible/acceptable behaviour" that acts as a constraint on how people exercise choice.[959] Likewise, the actions of many decision-makers can make it more difficult for some persons and groups to be free in their exercise of choice. The latter can be engendered through algorithmic acts of oppression. Persons and groups who are (systematically) excluded or underrepresented in data collection processes are made invisible and therefore more likely to be perceived as anomalous, deviant, or unknown by algorithmic decision-making systems. While the effects of each decision-instance might be limited, those harms can compound.

Algorithmic decision-makers might thus hold a position of dominance without their intending to.[960] Pettit for instance observed how unwilled domination "is likely to be exemplified in any relationships of inalienable asymmetrical power, as with the husband over the wife in a sexist culture, the employer over employees in an unregulated economy, the professor over students in a traditional college."[961] If a society does not provide people with sufficient (procedural) safeguards - such as data protection, consumer protection or non-discrimination laws, or the above-mentioned protective measures against oppression – to check decision-makers' ability to arbitrarily govern and control their choices, or does not enforce these safeguards, decision-makers are placed in a de facto dominating position.

---

[957] The term micro-domination is borrowed from O'Shea. See also: Danaher, 'Freedom in an Age of Algocracy', 266; John Danaher, 'Philosophical Disquisitions: Algorithmic Micro-Domination: Living with Algocracy', *Philosophical Disquisitions* (blog), 18 June 2018, https://philosophicaldisquisitions.blogspot.com/2018/06/algorithmic-micro-domination-living.html. O'Shea defined micro-domination as: "the capacity for decisions to be arbitrarily imposed on someone, which, individually, are too minor to be contested in a court or a tribunal, but which cumulatively have a major impact on their life." Tom O'Shea, 'Disability and Domination: Lessons from Republican Political Philosophy', *Journal of Applied Philosophy* 35, no. 1 (2018): 136, https://doi.org/10.1111/japp.12149.

[958] Danaher, 'Freedom in an Age of Algocracy', 266.

[959] Danaher, 266.

[960] Pettit, *On The People's Terms: A Republican Theory and Model of Democracy*, 62.

[961] Pettit, 62.

6.2 Algorithmic Stereotyping and Generalisations

Stereotypes concern the knowledge, beliefs, preconceptions and expectations persons hold about the characteristics, roles and attributes of groups of people, which in turn, are applied to the individual members of a group, often without regard to their actual situation.[962] While stereotypes can be framed positively and negatively, they maintain a standardised or generalised mental picture that reflects, and imposes onto people, an oversimplified or prejudiced attitude, assumption, belief or opinion.[963]

By attaching onto them a specific generalised world view, stereotypes can constrain a given group.[964] Their use can violate the (equal) human worth, dignity, and autonomy of those they capture. This might be so when stereotypes result in the subordination of a group, and its group members, and deny them the opportunity to participate fully in society and to express their identity.[965] Accordingly, stereotypes can limit people's right towards self-determination and self-development and become a tool for domination and oppression that helps maintain and reinforce existing power relationships.[966] Stereotypes can lead to wrongful prejudice (and vice versa) and stigmatise those marked out as others.[967] When a person or group faces generalised, prejudicial

---

[962] This definition predominantly builds upon the definition given by Alexandra Timmer, in reference to: Rebecca J. Cook and Simone Cusack, *Gender Stereotyping : Transnational Legal Perspectives* (Philadelphia PA: University of Pennsylvania Press, 2010), 9, https://www.worldcat.org/title/gender-stereotyping-transnational-legal-perspectives/oclc/320800041?page=citation. Timmer, *Strengthening the Equality Analysis of the European Court of Human Rights: The Potential of the Concepts of Stereotyping and Vulnerability.*, 45. However, it also takes cues from: Stephen J. Sherman et al., *Stereotype Development and Formation* (Oxford University Press, 2013), https://doi.org/10.1093/oxfordhb/9780199730018.013.0027.

[963] This particular definition draws from the Merriam Webster dictionary, which defines a stereotype as "something conforming to a fixed or general pattern, especially a standardized mental picture that is held in common by members of a group and that represents an oversimplified opinion, prejudiced attitude, or uncritical judgment." 'Definition of Stereotype', Merriam-Webster, 2022, https://www.merriam-webster.com/dictionary/stereotype.

[964] Whereas this Section approaches stereotypes in their negative connotation, under certain circumstances, stereotypes can be advantageous. For more information on the beneficial effects of stereotypes, see also: Sherman et al., *Stereotype Development and Formation*. See also: Schauer, *Profiles, Probabilities and Stereotypes*.

[965] Denise G Réaume, 'Discrimination and Dignity', *Louisiana Law Review* 63, no. 3 (2003): 53; Luke A Boso, 'Dignity, Inequality, and Stereotypes', *Washington Law Review* 92 (n.d.): 66. See also:

[966] Timmer, 'Toward an Anti-Stereotyping Approach for the European Court of Human Rights', 715.

[967] Solanke notes for instance that "Stigma begins with the deliberate social creation of a ' mark ' and results in individual discrimination." Solanke, *Discrimination as Stigma*, 37. Solanke provides a more in-depth look into the relationship and differences between stereotypes, prejudice and stigma. Solanke, chap. 1. Henri Tajfel, 'Stereotypes', *Race* 5, no. 2 (1963): 3, https://doi.org/10.1177/030639686300500201.

or erroneous beliefs, it undermines their position as social equals. They do not enter relationships on equal footing. Instead, they must fight of the generalised beliefs other may hold of them.[968] Stereotypes can thus impose a socio-relational status-harms onto others.[969] From this social inegalitarian position, stereotypes can be used to introduce, or argue in favour of, discriminatory or otherwise unfair differential treatment, which can be both socio-relational and socio-economic in character.[970]

Algorithmic differentiation grounds share a conceptual kinship with stereotypes in both appearance and function. First, group profiles derived from aggregates of (personal) data can often be redefined as generalised statements.[971] Second, like algorithmic differentiation grounds, stereotypes are a relational phenomenon: as a person becomes assigned as a member of one group, they are positioned against (members of) other groups.[972]

Algorithmic generalisations typically correspond with the idea of stereotypes as negative constraints in situations where they reinforce or maintain prejudice against historically disadvantaged or oppressed social groups. When Latanya Sweeney discovered how search queries using black-identifying names were more likely to return ads suggestive of an arrest record, algorithms reproduced the stereotype that black people would be more likely to engage

---

[968] See Chapters 4 and 5.

[969] While defined as a socio-relational status harm, stereotypes can arise out of socio-economic conditions (e.g., stereotypes regarding the poor) and have socio-economic consequences (e.g., less opportunities). Arguments against the use of stereotypes need not be defined in social egalitarian terms. For instance, as stereotypes are based on wrongful or erroneous beliefs, one could also argue that one should not use false information as a basis for decision-making. The unfair or unjustified nature of differential treatment lies not (only) in treating a group as inferior, but in using information that was false.

[970] See also Section 6.1.5 "Cultural imperialism."

[971] See also: Chapter 1, Section 1.2.2. "Group Level Knowledge." According to Schauer, stereotypes are little more than "generalizations in street clothes." Schauer, *Profiles, Probabilities and Stereotypes*, ix. Similarly, Hildebrandt ironically refers to stereotypes as a type of offline, everyday profile. Hildebrandt, 'Defining Profiling', 24.

[972] Other similarities can be found: First, like algorithmic categories, stereotypes have a structuring value: generalised statements reduce the complexity and nuance of society and therefore simplify people's interaction therein. Second, stereotypes allow persons to anticipate the world. In absence of specific knowledge regarding a thing or person beyond their membership of a group, they may be applied to predict and understand the behaviour of the group members the stereotype represents. By providing a simplification of the (social) world, a stereotype can be an efficient guide for social perception and behaviour. Third, like algorithmic categories, stereotypes do not need to be accurate in order for them to be applied. Rather, a loss of accuracy is made up for the efficiency they offer. Sherman et al., *Stereotype Development and Formation*. See for further information also: Walter Lippmann, *Public Opinion* (New York, Harcourt, Brace and Company, 1922), 90, http://archive.org/details/publicopinion00lippgoog; Tajfel, 'Stereotypes', 3; Sherman et al., *Stereotype Development and Formation*, 2.

in criminal activity.[973] Algorithms have been shown to reinforce and perpetuate existing prejudice and stereotypes vis-à-vis historically disadvantaged groups.[974]

In the next section, it will be explored to what extent algorithms can introduce novel stereotypes concerning non-socially salient and non-representational groups or aggregates. Whereas (negative) stereotypes constitute a clear socio-relational status harm, it will also be assessed whether more broadly speaking, the generalised treatment of decision-subjects could impose onto them socio-relational status harms.

6.2.1 Could (Novel) Algorithmic Stereotypes Arise?

Even though generalisations are relied upon by algorithmic decision-making systems, stereotypes are typically considered problematic where they have entered the shared consciousness of a given population. Could non-representational algorithmic differentiation grounds make this translation?

Where algorithmic differentiation does not yet represent or reflect clear and nefarious prejudice, their unique properties might render such a translation from the digital environment into the real world more difficult. Stereotypes are also more likely to form in relation to groups that have a higher level of uniformity.[975] non-socially salient, seemingly random, and non-tangible algorithmic differentiation grounds might lack the sense of (social) coherence that is needed for stereotypes to arise.[976] The algorithmic environment is notoriously opaque. People are typically unaware of the purposes for which, and the many ways through which, people can be ranked and categorised. If people are not aware of the relational dynamics imposed onto them by algorithms,

---

[973] Latanya Sweeney, 'Discrimination in Online Ad Delivery', 29 January 2013, https://arxiv.org/ftp/arxiv/papers/1301/1301.6822.pdf.

[974] See for instance: Noble, *Algorithms of Oppression*; Perez, *Invisible Women*.

[975] Sherman et al., *Stereotype Development and Formation*, 565.

[976] In literature, the coherence of a group has also been referred to as "entitativity". Sherman et al., *Stereotype Development and Formation*; Matthew T. Crawford, Steven J. Sherman, and David L. Hamilton, 'Perceived Entitativity, Stereotype Formation, and the Interchangeability of Group Members.', *Journal of Personality and Social Psychology* 83, no. 5 (2002): 1076–94; B. Lickel et al., 'Varieties of Groups and the Perception of Group Entitativity.', *Journal of Personality and Social Psychology* 78, no. 2 (2000): 223–46. See also on the difficulty of organizing people within the digital environment: Vedder, 'Medical Data, New Information Technologies and the Need for Normative Principles Other than Privacy Rules'; Anton H. Vedder, 'KDD, Privacy, Individuality, and Fairness', *Readings in CyberEthics, 2nd Edition. Jones and Bartlett, Sudbury, MA*, 2004, 462–70; Naudts, 'How Machine Learning Generates Unfair Inequalities and How Data Protection Instruments May Help in Mitigating Them'.

could they then formulate generalised assumptions regarding the groups that were created? Even if these groups would become known, people's response to them, might not be one that results in a lasting stereotype.[977] Tajfel believes that "when there is no emotional involvement in the use of the stereotype, when it represents no more than a moderately useful classifying device for lack of anything better to rely on, it is flexible and capable of change if information is received which flatly contradicts it or generally does not fit in."[978] The algorithmic environment might be too dynamic to introduce enduring stereotypes. For instance, in the case of adaptive learning processes, the attributes upon which decisions are built may change over time. In such a fast-evolving environment, groups might not be around long enough for them to become the subject of generalised assumptions.

At the same time, there may be reason to believe that in a highly interconnected and data-driven environment, new stereotypes could effectively form. For instance, certain attributes or groups that have proven to be relevant for decision-making could gain importance and permanence over time. Imagine an insurance company discovers that living in a rural area is strongly correlated to health and happiness. They transfer this information onto other institutions within the financial sector. As health and happiness signify a better life expectancy, credit institutions use these traits as a positive element when they assess people's capacity of repaying long-term loans. Over a prolonged period of time, this continued series of interconnected decisions results in a positive stereotype for people living in rural areas. Zarksy believes that the "wrongful assumption that a technological process exhibits a better capacity for rational reasoning than humans", may even act as a catalyst for rendering algorithmic generalisations more credible and acceptable.[979] When algorithmic differentiation grounds are able to enter into the minds of people as generalised

---

[977] In this regard, studies have shown how the social categorization of people into groups - even where that act has been unaffected by variables such as self-interest, interpersonal relations, conflict of interests or previous hostility – could be a sufficient condition for giving rise to adverse differential treatment and behaviour, against the "outgroup". In other words, excluding behaviour could be triggered by the mere division of people into groups. Henri Tajfel and Michael Billic, 'Familiarity and Categorization in Intergroup Behavior', *Journal of Experimental Social Psychology* 10, no. 2 (1 March 1974): 160, https://doi.org/10.1016/0022-1031(74)90064-X.

[978] Tajfel, 'Stereotypes', 12.

[979] Tal Z. Zarsky, 'An Analytic Challenge: Discrimination Theory in the Age of Predictive Analytics Cross-Cutting Issues of Law', *I/S: A Journal of Law and Policy for the Information Society*, no. 1 (2018 2017): 30. The tendency to ascribe more credence to machines has also been called automation bias, whereby humans overly rely on, and hence overlook the flaws of, automation. See for instance: Kate Goddard, Abdul Roudsari, and Jeremy C Wyatt, 'Automation Bias: A Systematic Review of Frequency, Effect Mediators, and Mitigators', *Journal of the American Medical Informatics Association : JAMIA* 19, no. 1 (2012): 121–27, https://doi.org/10.1136/amiajnl-2011-000089.

beliefs or assumptions regarding a specific group, and their group members, such could materialise into socio-relational harms similar to those currently associated with stereotypes.

6.2.2 The Problematic Nature of Generalisations: Deindividualisation and Misrepresentation.

Stereotypes and generalisations are incomplete sets of knowledge regarding a specific group and their group members. In principle, these generalised beliefs should be discarded, and people should modify their views as soon as they learn more about the actual behaviours of those groups.[980] Unfortunately, stereotypes and generalisations are not always applied with such a level of care. Even when they are proven to be wrongful, racist, sexist or otherwise unjustified, stereotypes can still persist. Whereas the socio-relational harms of wrongful stereotypes and generalisations are now known, in this Section, I address the question whether generalisations in general could constitute a socio-relational status-harm for the decision-subjects to which they apply.

DEINDIVIDUALISATION - According to Vedder, algorithmic generalisation processes can give rise to the deindividualisation of the individual: "persons are judged and treated more and more as members of a group (i.e., the reference group that makes up the data or information subject), rather than as individuals with their own characteristics and merits."[981] Algorithms therefore fail to treat individuals *as* individuals.

The Finnish Non-Discrimination and Equality tribunal used a similar argument against the deployment of algorithmic statistical models for credit scoring decisions. The case concerned the denial of a credit application solely based on formal and abstract credit data based on the behaviour of others. While the model was based on prohibited grounds, the tribunal also observed how the credit scoring system ignored individualised information regarding the applicant's credit behaviour, even though these factors would have favoured credit provision. The Tribunal considered that a sole reliance on statistical models and without performing an individual assessment of the applicant's financial standing was disproportionate and not acceptable under Finnish non-discrimination law. [982]

---

[980] Lippmann, *Public Opinion*, 90–91.

[981] Vedder, 'KDD', 277.

[982] Lorenz Matzat and Minna Ruckenstein, 'Finnish Credit Score Ruling Raises Questions about Discrimination and How to Avoid It', *AlgorithmWatch* (blog), 21 November 2018, https://algorithmwatch.org/en/finnish-credit-score-

Eidelson has observed how a failure to treat people as individuals need not imply that the use of any statistical or generalised way of approaching individuals should be prohibited. Rather, in making use of such techniques, people should always be treated in a way that pays close attention "to their own roles in determining who they are, and [respect] their capacities to make choices for themselves as autonomous agents."[983] Under this account, generalisations should not undermine a person's fundamental interest towards self-determination. The use of generalisations does indeed appear particularly problematic when they are used to predict the actions and behaviour of persons without them having any recourse to address those decisions based on their actual or individualised information. For one, rather than take into account the actual autonomous choices a person has made, generalisations assume a person's past behaviour and actions.[984] Even in the case these generalisations do correspond to a person's past (autonomous) actions and behaviour, they might still restrain a person's self-determination towards the future by imposing onto them a form of data determinism.[985] A person is tied to their past and denied full authorship over their future actions. In both cases, actual choice and behaviour are discarded in favour of predicted choice and behaviour. [986] This failure to treat people as individuals would be sufficient to render a specific decision unfair or unjustified. The inequalities a person incurs as a result of those

---

ruling-raises-questions-about-discrimination-and-how-to-avoid-it/.; National Non-Discrimination and Equality Tribunal of Finland (21 March 2018, 216/2017), translation available via: https://www.yvtltk.fi/material/attachments/ytaltk/tapausselosteet/45LI2c6dD/YVTltk-tapausseloste-_21.3.2018-luotto-moniperusteinen_syrjinta-S-en_2.pdf

[983] Benjamin Eidelson, 'Treating People as Individuals', in *Philosophical Foundations of Discrimination Law* (Oxford: Oxford University Press, 2013), 227, https://doi.org/10.1093/acprof:oso/9780199664313.003.0011.

[984] Binns warns against the use of arguments of this kind: "What appear to be criticisms of generalization in general(!), may in fact boil down to criticisms of insufficiently precise means of generalization." Binns, 'Fairness in Machine Learning: Lessons from Political Philosophy', 5. H

[985] Data determinism" was coined by Ramirez. Citron and Pasquale refer to this phenomenon as "arbitrariness-by-algorithm." Edith Ramirez, 'The Privacy Challenges of Big Data: A View from the Lifeguard's Chair', in *Keynote Address* (Technology Policy Institute Aspen Forum, Aspen, 2013), https://www.ftc.gov/sites/default/files/documents/public_statements/privacy-challenges-big-data-view-lifeguard%E2%80%99s-chair/130819bigdataaspen.pdf. The notion "; Daniel Keats Citron and Frank Pasquale, 'The Scored Society: Due Process for Automated Predictions', *Wash. L. Rev.* 89, no. 1 (2014): 1–33; Zarsky, 'Understanding Discrimination in the Scored Society'.

[986] A person may have a steady job, good pay grade and healthy lifestyle. These conditions strongly suggest a person's capability to repay a debt. Yet, an algorithmically guided decision-making process may still deny the loan application because of the person's handwriting corresponds to the handwriting of various lenders who ended up being non-trustworthy. Whereas both statements can be true to some extent, they lead to opposite conclusions. Vedder, 'KDD'; Vedder and Naudts, 'Accountability for the Use of Algorithms in a Big Data Environment', 210.

decisions, such as lesser socio-economic prospects due to the denial of the credit application, would then be unjustified as well.

MISREPRESENTATION - At the same time, the incontestable use of generalised statements is a type of misrecognition or misrepresentation. The persons captured by the generalisation are defined and perceived on the basis of their (alleged) group membership, rather than their individual actions and behaviour. Even though the applicant would have been eligible for credit, they were misrecognised and misrepresented by the system as non-trustworthy.[987] Without any possibility for recourse, this type of misrecognition can lead to a process of alienation. According to Schuppert, alienation occurs when a person no longer sees themselves as a legitimate source of claims and reasons.[988] Due to the experience of misrecognition, as well as their powerlessness and voicelessness to act against the decision, their capacity towards self-development and self-determination is undermined: they do not have the ability to advance their own proper reasons and judgments to counteract the generalised belief imposed onto them.[989]

Because their deindividualizing and misrepresentational character actively interfere with people's right towards self-determination and self-development, the use of algorithmic generalisations constitutes an algorithmic act of domination and oppression. This will be the case when decision-makers deny decision-subjects the option to exercise control over, and participate in, the conditions of their actions and behaviours, and deprive them of the opportunity to express their point of view and experiences in recognised ways.

Following the aforementioned, decision-makers' current reliance on generalised statements need not be prohibited. If decision-makers need to take seriously the idea that decision-subjects should remain their own self-authors however, they should give them the opportunity to contest and argue against these generalisations and the outcomes they produce. A start to such an approach towards the regulation of algorithmic decision-making can be found in the GDPR. For example, Article 22 GDPR gives data subjects (under certain circumstances) the right not to be subject to

---

[987] When the generalised statement is proven to be wrong, misrepresentation not only occurs on the individual level, but also on the group level.

[988] Schuppert, 'Non-Domination, Non-Alienation and Social Equality', 447.

[989] Schuppert, 447. See also: Garrau and Laborde, 'Relational Equality, Non-Domination, and Vulnerability'. Sylvie Delacroix and Michael Veale, 'Smart Technologies and Our Sense of Self: Going beyond Epistemic Counter-Profiling', *Life and the Law in the Era of Data-Driven Agency*, 31 January 2020, https://www.elgaronline.com/view/edcoll/9781788971997/9781788971997.00011.xml; Andrea Sangiovanni, *Humanity without Dignity: Moral Equality, Respect, and Human Rights.* (Harvard University Press, 2017).

decisions based solely on automated processing, including profiling, which produce legal effects concerning, or similarly significantly affect, them.[990] When such decisions do take place, data subjects ought to be provided with adequate protection, such as the right to obtain human intervention, the ability to express their point of view or the possibility to contest the decision.[991] Furthermore, in these cases the GDPR also expects data controllers to provide data subjects with meaningful information regarding the logic involved, as well as the significance and the envisaged consequences of such processing for the data subject.[992]

Without measures that enable decision-subjects to individualise and represent themselves against the outcomes produced by generalised statements, decision-makers will fail to treat and recognise them as self-authors. This would undermine decision-subjects in exercising their fundamental interest to self-development and self-determination, and doing so, would impose onto them a socio-egalitarian status-harm.

---

[990] Article 22, § 1 GDPR. Much has been said about this provision, the conditions under which it applies and its deficits in practice. See for example: Bryce Goodman and Seth Flaxman, 'EU Regulations on Algorithmic Decision-Making and a "Right to Explanation"', *2016 ICML Workshop on Human Interpretability in Machine Learning (WHI 2016)*, 2016, 1–9; Gianclaudio Malgieri and Giovanni Comandé, 'Why a Right to Legibility of Automated Decision-Making Exists in the General Data Protection Regulation', *International Data Privacy Law* 7, no. 4 (1 November 2017): 243–65, https://doi.org/10.1093/idpl/ipx019; Gianclaudio Malgieri, 'Automated Decision-Making in the EU Member States: The Right to Explanation and Other "Suitable Safeguards" in the National Legislations', *Computer Law & Security Review* 35, no. 5 (1 October 2019): 105327, https://doi.org/10.1016/j.clsr.2019.05.002; Lilian Edwards and Michael Veale, 'Slave to the Algorithm? Why a "right to an Explanation" Is Probably Not the Remedy You Are Looking For.', *Duke Law & Technology Review* 16, no. 1 (23 May 2017): 18–83, https://doi.org/Edwards, Lilian and Veale, Michael, Slave to the Algorithm? Why a 'Right to an Explanation' Is Probably Not the Remedy You Are Looking For (May 23, 2017). 16 Duke Law & Technology Review 18 (2017). Available at SSRN: https://ssrn.com/abstract=2972855 or http://dx.doi.org/10.2139/ssrn.2972855; Sandra Wachter, Brent Mittelstadt, and Luciano Floridi, 'Why a Right to Explanation of Automated Decision-Making Does Not Exist in the General Data Protection Regulation', *International Data Privacy Law* 7, no. 2 (2017): 76–99; Andrew D. Selbst and Julia Powles, 'Meaningful Information and the Right to Explanation', *International Data Privacy Law* 7, no. 4 (1 November 2017): 233–42, https://doi.org/10.1093/idpl/ipx022; Article 29 Data Protection Working Party, 'Guidelines on Automated Individual Decision-Makinng and Profiling for the Purposes of Regulation 2016/79' (Article 29 Data Protection Working Party, 2018); Naudts, Dewitte, and Ausloos, 'Meaningful Transparency through Data Rights: A Multidimensional Analysis'; Margot E. Kaminski, 'The Right to Explanation, Explained', *Berkeley Technology Law Journal* 34 (2019): 189–218, https://doi.org/10.15779/Z38TD9N83H.

[991] Article 22, § 3 GDPR.

[992] Articles 13, §2 (f); 14, §2 (g) and 15, §1(h) GDPR.

6.3 Conclusion

In this Chapter, I examined the function socio-relational conceptualisations of equality can have for the evaluation and regulation of algorithmically guided decision-making structures and acts. Social equality shifts the evaluative focus of attention towards the relationships (vis-à-vis themselves, others, groups, private corporations, or governments) decision-subjects stand in within the algorithmic environment. Making use of the notions of oppression and domination, as well as deindividualisation and misrepresentation, it was demonstrated how algorithmic structures and decision-making acts can impose a wide variety of socio-relational egalitarian harms onto decision-subjects that are inimical to their status as free, equal, and valued members of society.

OPPRESSION AND DOMINATION – Algorithmic oppression refers to the capacity for decision-makers to i) exploit, ii) marginalise, iii) render powerless decision-subjects (or take advantage of their powerlessness), iv) impose onto them (subjective) world views in ways that render their lived experiences unheard or invisible (cultural imperialism), or v) expose decision-subjects to (digital) forms of violence. Algorithmic domination can be understood as the capacity for decision-makers to interfere in decision-subjects' valuable options and choices in a manner that is arbitrary, unchecked, or uncontrolled. Oppression and domination stand opposed to self-development (the capacity to learn, develop, exercise, and express skills, capacities, and experiences in socially meaningful and recognised ways) and self-determination (the capacity to exercise control over, and participate in determining, their own options, choices, and the conditions of their actions). When a person or group are the subject of domination or oppression they suffer a status-inequality; it placed onto them a vulnerability-marker that signals their unequal social standing. Through the actions of others, they are made voiceless and invisible. In order to protect their fundamental interest to self-determination and self-development, decision-subjects should be equal in their enjoyment to navigate the algorithmic environment free from domination and oppression.

Decision-makers should exercise their powers over decision-subjects responsibly. On the one hand, decision-makers should refrain from being oppressive or dominating towards their decision-subjects. Otherwise, they would take advantage of the vulnerable position decision-subjects find themselves in. On the other hand, given the power they hold, decision-makers should make sure that their actions or decisions do not intensify a person or group's risk to become the subject of domination and oppression by others. Indeed, depending upon the context and conditions under which a decision is made, algorithmic differentiation acts, and the inequalities

they generate, can maintain, decrease, or increase a person or group's vulnerability status vis-à-vis others.

STEREOTYPING, GENERALISATIONS AND DEINDIVIDUALISATION – Algorithmic differentiation grounds share a conceptual relationship with stereotypes. Like stereotypes, they can be reformulated as a generalised statement that places person in a relationship with and against others. Stereotypes hinder persons in their capacity to enter social relationships on equal footing. Instead, their initial position is pre-determined by the erroneous or false assumptions that have been attached to them. In the case the generalised assumptions that are reflected within algorithmic differentiation grounds find entry into the minds of people, they too could constrain persons and groups like stereotypes do.

Generalised statements do not need enter the shared consciousness of people in order to be considered problematic. The argument was made that algorithmic decision-makers' reliance on generalising differentiation grounds could undermine decision-makers' social standing if the use of these grounds cannot be contested. Generalisations deindividualise decision-subjects. They are judged on the basis of assumed or predicted, rather than individualised actions, behaviours, and characteristics. Decisions based upon generalisations fail to consider the autonomous choices decision-subjects have made or can make if there are no mechanisms for individual redress. Second, this process could also lead to the misrepresentation and misrecognition of decision-subjects, who are entirely defined on the basis of their alleged group membership and are treated in correspondence thereto. Put differently, the incontestable use of generalisations fails to recognise decision-subjects as self-authenticating self-authors. Due to the autonomy-limiting predictions of decision-makers, combined with the powerlessness and voicelessness to act against them, decision-subjects are denied the opportunity to govern the conditions of their own actions and advance their own reasons and judgments in socially recognised ways. Both stereotypes and generalisations therefore have the capacity to undermine a person or group's right towards self-determination and development, and as a consequence, their use could constitute an algorithmic act of oppression and domination.

POSITIVE MEASURES – Regulators should guard and protect decision-subjects from oppression and domination, including unjustifiable and incontestable generalisations. Having articulated the socio-egalitarian status harms decision-subjects can suffer, and the conditions under which those harms can arise, several regulatory recommendations could be formulated with a view to strengthen decision-subjects in their deliberate navigation of the algorithmic environment. Mechanisms of (increased) accountability, transparency, participation, and contestation were

identified as valuable resources that can guard decision-subjects (and algorithmic groups) against unjustifiable interferences in their rights towards self-development and self-determination. These tools can give decision-subjects or other interest groups an opportunity to break down asymmetries in power and regain democratic control over their living environment and the structures and decisions that mediate and shape their lives. At the same time, regulators should continue to encourage decision-makers to reflect upon their position of privilege. The latter should map and monitor how their actions, including the socio-relational and socio-economic inequalities they produce, could affect the vulnerability status of decision-subjects.

7 Distributive Conceptualisations of Equality and Algorithmically Guided Decision-Making

7.1 Introduction

Following a distributive conceptualisation of equality, policy discussions centre around the allocation of burdens and benefits across people. Departing from a specific fundamental interest that people share, a distributive framework defines a specific set of (non)material conditions or justice-relevant goods that people should have (roughly) equal access to in order to realise those interests.[993] The fairness of inequalities can be evaluated in light of people's outcomes in their distributive shares. A distributive account formulates rules on how a given society can achieve the best possible distribution, and the actions of actors and institutions can be assessed in terms of their contribution to these distributions.[994]

In the algorithmic environment, distributive conceptualisations of equality can be used to evaluate the deployment of algorithms in their function of producers of outcomes/distributors of opportunities, risks, and benefits among people. From a regulatory perspective, algorithmic distributive justice demands that the algorithmic processes, which mediate and co-shape the basic structure of society, do not interfere with the just distribution of justice-relevant goods.[995] When decision-makers impose differentiation onto decision-subjects, and as a consequence thereof, the desired or just distribution of justice-relevant goods is no longer maintained, the resulting (in)equalities constitute an allocative or distributive harm. Conversely, a decision-maker's actions

---

[993] These justice-relevant goods represent a specific metric of justice, such as *equality* of welfare, *equality* of resources, or as discussed within this Chapter, *equality* of primary goods and capability. Furthermore, distributive equality does not mandate strict equality and deviations from people's equal holdings of certain justice-relevant goods could be justified. For example, people might still be held responsible for the individual choices they make that affect their holding of a specific good, or there might be other interests that compete with and override equality. Schemmel, 'Distributive and Relational Equality', 126–27; Pogge, 'Relational Conceptions of Justice', 2004; Voigt, 'Distributive Equality, Relational Equality and Preferences about Higher Education'; Kristin Voigt, 'Relational Equality and the Expressive Dimension of State Action', *Social Theory and Practice* 44, no. 3 (2018): 437–67. Sen, 'Equality of What?'; Cohen, 'On the Currency of Egalitarian Justice'.

[994] Schemmel, 'Distributive and Relational Equality', 127. In reference to Pogge, 'Three Problems with Contractarian-Consequentialist Ways of Assessing Social Institutions*', 242–46. and Thomas W. Pogge, 'Relational Conceptions of Justice: Responsibilities for Health Outcomes', *Veritas* 46, no. 1 (1 January 2001): 51–75.

[995] Rawls, *A Theory of Justice*, 1971, 7.

are justified when the (in)equalities they generate do contribute to the desired distributive state of affairs.

It is worth noting that when we purely focus on people's outcomes, like some distributive conceptualisations may suggest we do, we may obscure how people's social and economic position is shaped by relational processes and structures. We should not only pay attention to what people have access to, but also how people are, and should be, treated in the process. In other words, we need to closely monitor the relational dynamics of equality. Still, in order to realise the socio-relational ideal of a community of equals, where citizens stand on equal social footing with one another, people do need access to certain opportunities, benefits or other (non)material goods so they can successfully exercise their right to self-determination and self-development. Informed by a socio-relational ideal of justice, distributive choices can therefore complement the former by highlighting those domains of life, as well as (non)material goods people need access to in order to stand in relationships of equality.[996] While insufficient in themselves to secure social equality, the identification of justice-relevant goods helps regulators to point out those areas of life where algorithmic decision-makers hold power, and where their decisions could have a (greater) capacity to render decision-subjects more vulnerable on a socio-relational level.[997]

This Chapter examines two approaches towards distributive justice in an effort to evaluate the function of algorithms as producers of outcomes/distributors of opportunities, benefits, and burdens: equality of primary goods (John Rawls, Section 7.2), and equality of capability (Martha Nussbaum, Section 7.3).[998] The following Sections are structured in light of the aforementioned description of the distributive paradigm. In a first step, I identify the specific fundamental interest people are said to share. In a second step, I clarify how those interests can be secured through the distribution of primary goods and central capabilities, respectively. This analysis makes it possible to further strengthen our understanding of the egalitarian harms algorithmic decision-makers risk to impose and the safeguards that could be considered and put forward in order to protect decision-subjects.

---

[996] While both accounts emphasise the distributional, rather than the relational nature of justice, they nonetheless recognise how societies are built upon ideals of mutual respect and recognition, and the fundamental interests they identify, can be linked to values (similar to) self-determination and self-development.

[997] Scheffler, 'What Is Egalitarianism?', 2003; Anderson, 'What Is the Point of Equality?' Scheffler, 'The Practice of Equality'.

[998] Young, *Inclusion and Democracy*, 31–32.

7.2 Equality of Primary Goods

According to Rawls, a society must define a set of principles through which the shared and competing interests of citizens can be arranged through the designation of rights, duties and a definition of an appropriate distribution of benefits and burdens.[999] In his theory of justice, Rawls' conceptualises justice as fairness.[1000] Rawls believes that a suitable conception of justice can be agreed upon through mutual agreement reached under fair conditions between individuals that are conceived as free and equal moral persons. The "fairness of the circumstances under which agreement is reached, transfers to the fairness of the principles agreed to."[1001] Rawls' principles of justice are both liberal and egalitarian: while they affirm the importance and priority of basic rights and liberties, they also show a deep concern for entrenched social and socio-economic inequalities. The principles are distributive as they govern the allocation of so-called primary goods: all-purpose goods that free and equal citizens require in order to have access to a complete life.[1002]

[999] Rawls, *A Theory of Justice*, 1971, 4.

[1000] Rawls, *A Theory of Justice*, 1971; Rawls, *Justice as Fairness: A Restatement*.

[1001] 'A Kantian Conception of Equality by John Rawls '43 *50', accessed 17 August 2020, https://princetonindependent.com/issue01.03/item10c.html. This fairness is guaranteed as a result of the fair conditions under which people are able to reach an agreement regarding the principles of justice. They are the outcome, of an original agreement in a suitably defined initial situation. The participants of this initial situation - or the original position as it is referred to within a Theory of Justice - decide upon the principles that will govern their association. While the participants are rational and in the pursuit of their own interests, they are also each other's equals. What follows is a situation fairly common in social theory; a simplified situation in which "rational individuals with certain ends and related to each other in certain ways are to choose among various courses of action in view of their knowledge of the circumstances." Furthermore, the parties that partake in this hypothetical original position are situated behind a veil of ignorance. No participant knows their place in society, class, or status, nor what natural talents, such as intelligence or strength, they might possess. A participant does not know the details of their life, nor their psyche. They also do not know the specific type of society they belong too. For example, whether that society would be cultured, in recession or in an economic boom. Rawls, *A Theory of Justice*, 1971, chaps 3, The Original Position, and p. 119.

[1002] Rawls, *Justice as Fairness: A Restatement*, 58.

7.2.1 People as Free and Equal: Two Moral Powers

The principles of justice do not seek to promote or realise a specific conception of the good.[1003] Instead, they look to structure society in such a way that citizens are given the means they need to realise their fundamental interests as free and equal persons.[1004]

PEOPLE AS EQUALS – According to Rawls, persons have two so-called moral powers.[1005] On the one hand, people have the capacity for a sense of justice. This moral power refers to people's capacity "to understand, to apply, and to act from (and not merely in accordance with) the principles of political justice that specify the fair terms of social cooperation."[1006] A person's sense of justice finds expression in their reasonable and "regulative desire to act upon certain principles."[1007] On the other hand, people have a capacity to form, revise and (rationally) pursue their own conception of the good.[1008] People are equal in their (shared) capacity or potentiality to develop and exercise their so-called moral powers.[1009]

[1003] Rawls, *A Theory of Justice*, 1971, 560. To take the latter as the point of departure in structuring society, is, according to Rawls, to have one's order wrong.

[1004] Rawls, 560. Leif Wenar, 'John Rawls', in *The Stanford Encyclopedia of Philosophy*, ed. Edward N. Zalta, Spring 2017 (Metaphysics Research Lab, Stanford University, 2017), https://plato.stanford.edu/archives/spr2017/entries/rawls/.

[1005] This idea of the person is both normative and political, not metaphysical or psychological. Rawls, *Justice as Fairness: A Restatement*, 18–19.

[1006] Rawls, 18–19. Wenar, 'John Rawls'.

[1007] Rawls, *A Theory of Justice*, 1971, 561.

[1008] According to Rawls, happiness was too vague and abstract a concept to adequately account for a diversity of world views, and too difficult to measure in order to be practicable. Initially, Rawls sought refuge to the Aristotelian principles and rational life plans, but later revised his approach to develop a criterion of justice that was not necessarily informed by all needs and interests of human beings, but rather by their needs and interests as citizens of a free democratic society. See also: Thomas Pogge, *John Rawls: His Life and Theory of Justice*, trans. Michelle Kosch (New York: Oxford University Press, 2007), 53–55.

[1009] Rawls, *A Theory of Justice*, 1971, para. 77 (The Basis of Equality); Rawls, *Justice as Fairness: A Restatement*, 20. Similarly, people are equal in the sense that every person has, and views themselves as having, the right towards equal respect and consideration in choosing the principles that will govern society. In this regard, Guyer observes how "two moral powers are the basis of the conception of ourselves as free and equal persons, and the commitment to them is the basis of wanting to express our conception of ourselves as free and equal persons."[1009] John Rawls, 'A Kantian Conception of Equality', *Cambridge Review*, 1975, 94–99; Paul Guyer, 'Principles of Justice, Primary Goods and Categories of Right: Rawls and Kant', *Kantian Review* 23, no. 4 (December 2018): 581–613, https://doi.org/10.1017/S1369415418000419.

PEOPLE AS FREE – People are free in two key ways.[1010] First, people are free when they are not tied to, nor identified with, a specific conception of the good, but capable to form, pursue, and if need be revise what they themselves consider valuable and worthwhile.[1011] Second, people are free when they regard themselves as "self-authenticating sources of valid claims": they are entitled to make claims in an effort to advance their own conception of the good and they have confidence that these claims can be heard.[1012]

This conceptualisation of freedom roughly corresponds to the notions of self-determination (being able to exercise control and participate in the conditions of your options and actions) and self-development (being able to exercise and express your skills and experiences in socially recognised settings and to express their views on social life).[1013] Each person has an inviolable capacity for freedom, self-determination and self-development: "a moral person is a subject with ends [they have] chosen, and [their] fundamental preference is for conditions that enable [them] to frame a mode of life that expresses [their] nature as free and equal rational being as fully as circumstances permit."[1014]

THE PRIMARY GOODS - The fundamental interest of free and equal persons lies in their being able to develop and exercise their moral powers, and to be free in their effective pursuit of their own particular conception of the good.[1015] The primary goods are considered essential for people to do so, and are therefore what "free and equal persons *need* as citizens."[1016] They are the all-

---

[1010] John Rawls, *Political Liberalism* (New York, Chichester, West Sussex: Columbia University Press, 1993), 29–35.

[1011] Rawls, 30. Under the assumption that people can control and revise their wants and desires, they also hold responsibility for doing so. Rawls, 'A Kantian Conception of Equality'.

[1012] In addition to people's being free in the abovementioned sense, it has also been argued that people are free in the sense that they are capable of taking responsibility for their own ends. Rawls, *Political Liberalism*, 33. Wenar, 'John Rawls'. Rawls, *Political Liberalism*, 32. Rawls, *Justice as Fairness: A Restatement*, 23.

[1013] See Chapter 6: Section 6.1 Algorithmic Oppression and Domination. Young, *Justice and the Politics of Difference*, 38.

[1014] Rawls, *A Theory of Justice*, 1971, 561. Alexander, John M., 'Capability Egalitarianism and Moral Selfhood', *Ethical Perspectives*, no. 1 (2003): 3, https://doi.org/10.2143/EP.10.1.503868. A person's conception of the good finds expression in their "rational plan of life": an ordered family of final ends and aims which specifies a person's conception of what is of value in human life or, alternatively, of what is regarded as a fully worthwhile life." For their conceptualisation or interpretation of the good, people can draw inspiration from (among others) religion, philosophy or moral doctrines. Rawls, *A Theory of Justice*, 1971, 561. Rawls, *Justice as Fairness: A Restatement*, 19.

[1015] Rawls, *Justice as Fairness: A Restatement*, 85.

[1016] John Rawls, 'The Basic Liberties and Their Priority', ed. Sterling M Mcmurrin, The Tanner Lectures on Human Values, 1981, 89. Rawls, *Justice as Fairness: A Restatement*, 60. Rawls, 'A Kantian Conception of Equality'; John Rawls, 'A Well-Ordered Society', in *Philosophy, Politics and Society*, ed. Peter Laslett and James S. Fishkin, vol. 5 (Oxford, UK: Blackwell Publishers, 1979), 6–20.

purpose means through which a "sufficiently wide range of final ends" can be realised.[1017] People therefore also have a fundamental interest to maximise their access to these goods: the higher a person's share of primary goods, the better their position.[1018] The primary goods are the basic rights and liberties[1019], freedom of movement and free choice of occupation[1020], the powers and prerogatives of offices and positions of responsibility[1021], income and wealth[1022], and the social

[1017] Rawls, 'The Basic Liberties and Their Priority', 22. Even though Rawls admits that one should take cues from human needs and abilities, the relationships people want to seek out, or how "rational plans of life" are typically structured in order to settle on a list of primary goods, these general facts and requirements of social life are insufficient. Instead, the primary goods are" objective" and chosen primarily in light of Rawls' idealised, political, and normative conception of the person. The primary goods are therefore also not defined as the result of some historical, social, or psychological survey, nor do they follow a "comprehensive moral doctrine." Likewise, their good is not assessed "in light of some personal ideal." The lack of objectivity clarifies Rawls' objection against the capability approach, which he perceived as endorsing a comprehensive moral view. Rawls, *Justice as Fairness: A Restatement*, 60. Ingrid Robeyns and Harry Brighouse, 'Introduction: Social Primary Goods and Capabilities as Metrics of Justice', in *Measuring Justice*, ed. Harry Brighouse and Ingrid Robeyns (Cambridge: Cambridge University Press, 2010), 5, https://doi.org/10.1017/CBO9780511810916.001. Rawls, *Justice as Fairness: A Restatement*, 58–59.

[1018] The primary goods are social, rather than natural in nature. As a consequence, people's well-being is assessed on the basis of their holding of primary goods, rather than the natural abilities and talents, such as their intelligence or stamina, they have. While people's holdings of "natural goods" also affect their well-being, these goods cannot be directly controlled by the institutions that must provide justice. *John Rawls: His Life and Theory of Justice*; Pogge, *John Rawls: His Life and Theory of Justice*, 74. Rawls, *A Theory of Justice*, 1971, 62. Amartya Sen, *The Idea of Justice* (Belknap Press: Harvard University Press, 2009); Martha C. Nussbaum, *Frontiers of Justice: Disability, Nationality, Species Membership*, The Tanner Lectures on Human Values (Cambridge, Massachusetts, London: Belknap Press: Harvard University Press, 2006). Wenar, 'John Rawls'.

[1019] The basic rights and liberties are indispensable for the fully and informed exercise of moral powers and protect a wide range of conceptions of the good. Rawls, 'The Basic Liberties and Their Priority'; Rawls, *Justice as Fairness: A Restatement*, 44. See also Section 7.2.3 The Basic Liberties

[1020] Freedom of movement and free choice of occupation must be seen against a background of diverse opportunities. The freedom these primary goods offer, enable people to pursue different ends and, if necessary, make changes in the plans they have set out for themselves.

[1021] The powers and prerogatives of offices and positions of responsibility give "scope to various self-governing and social capacities of the self." Rawls, 'The Basic Liberties and Their Priority', 23.

[1022] Income and wealth serve more generally as the type of "all-purpose means". Income and wealth can be (in)directly used to achieve and advance a wide variety of ends. The primary goods are not always as well defined. For example, with regard to wealth, Rawls states the following: "This [wealth] is not an easy concept to define, but I mean to use it roughly in the sense understood by economists. Thus, wealth consists of (legal) command over exchangeable means for satisfying human needs and interests. Items that we own, such as food and land, buildings and machines, are wealth; so too are rights to use or to receive or in any way to derive benefit from such items-for example, shares in private or public companies, or rights of access to libraries, museums, and other public facilities, rights to various kinds of personal services, and so on." John Rawls, 'Fairness to Goodness', *The Philosophical Review* 84, no. 4 (October 1975): 540, https://doi.org/10.2307/2183853.

bases of self-respect[1023]. They constitute the main metric through which a given state of affairs, and the (in)equalities between people, can be evaluated.[1024]

7.2.2 The Principles of Justice

Primary goods must be distributed in correspondence with the principles of justice. The first principle stipulates that "each person has the same indefeasible claim to a fully adequate scheme of equal basic liberties, which scheme is compatible with the same scheme of liberties for all."[1025] The second principle postulates that for social and economic inequalities to be justified they must satisfy two conditions. The first condition requires that socio-economic inequalities ought to be attached to offices and positions open to all under conditions of fair equality of opportunity. The second condition is referred to as the difference principle. This principle requires that socio-economic inequalities must produce the greatest benefits for the least-advantaged members of society.[1026]

EQUALITY - The aforementioned principles are informed by Rawls' intuitive understanding that people "born into different positions have different expectations of life determined, in part, by the political system as well as by economic and social circumstances."[1027] These differing starting positions favour certain starting positions over others, and these create "especially deep inequalities": they are pervasive, systemic, undeserved, and unmerited. Rawls distinguishes between three broad categories of contingencies as having a profound impact on the inequalities

---

[1023] The social bases of self-respect give citizens a sense of self-worth and the capacity to exercise their moral powers and pursue their ends with self-confidence. Rawls, *Justice as Fairness: A Restatement*, 60. See also Section 7.2.3 "The Social Bases of Self-Respect."

[1024] Other metrics may include "utility", "resources"," welfare" or as will be discussed in the next section: "capabilities."

[1025] See also Section 7.2.3 "The Basic Liberties" In the original Theory of Justice, Rawls furthermore noted that liberty can only be restricted for the sake of liberty. The latter comprises only two instances: "(a) a less extensive liberty must strengthen the total system of liberty shared by all; (b) a less than equal liberty must be acceptable to those with the lesser liberty." Rawls, *A Theory of Justice*, 1971, 302.

[1026] See also Section 7.2.3. The Difference Principle. It can also be noted that the principles have been given a specific (lexical) ordering. Above everything else, people should be guaranteed the basic liberties. Once those have been guaranteed, the goal is to ensure fair equality of opportunity. Finally, within a setting where both basic equal liberties and fair equality of opportunity has been guaranteed, the least-advantaged members of society should benefit from the social and economic inequalities of society. Rawls, *Justice as Fairness: A Restatement*, 43. On the lexical ordering of the principles, see also: Rawls, *A Theory of Justice*, 1971, 302–3.

[1027] Rawls, *A Theory of Justice*, 1971, 7. Wenar, 'John Rawls'.

that exist between people in terms of their life's prospects[1028]: a) people's social class of origin; b) their native endowments; and their opportunities to develop these endowments as affected by their social class of origin and c) their fortune or luck, both good and bad, over the course of their life. These contingencies are morally arbitrary, and the principles of justice are designed to counterbalance the inequalities in people's prospects that arise from them.[1029]

Whereas the previous Chapter demonstrated how algorithms can violate people's fundamental interest to self-determination and self-development through algorithmic acts of oppression and domination, including unjustified generalisations, under a Rawlsian conception of equality, the detrimental impact that algorithmic acts have on these interests can now be evaluated through a distributional lens: do algorithmic structures limit, or rather contribute to, people's access to, and enjoyment of the primary goods (in accordance to the principles of justice), and can the inequalities algorithmic decision-makers introduce between decision-subjects be justified in view of this desired distribution?[1030]

Equal Basic Liberties

According to the first principle, each person should have the same access to a set of basic rights and liberties.[1031] The basic liberties include at a minimum: "freedom of thought and liberty of conscience; the political liberties and freedom of association, as well as the freedoms specified by the liberty and integrity of the person[1032]; and finally, the rights and liberties covered by the rule

[1028] Rawls, *Justice as Fairness: A Restatement*, 55–56.

[1029] Rawls, *A Theory of Justice*, 1971, 72.

[1030] Alexander, John M., 'Capability Egalitarianism and Moral Selfhood', 3. Generally phrased: "All social primary goods – liberty and opportunity, income and wealth, and the bases of self-respect – are to be distributed equally unless an unequal distribution of any or all of these goods is to the advantage of the least favoured." Rawls, *Justice as Fairness: A Restatement*, 57. Rawls, *A Theory of Justice*, 1971, 303.

[1031] The basic liberties are the same for every person. The worth or usefulness of liberty might not be the same for everyone however. Some people will have more income and wealth and hence more opportunities to pursue their vision of the good. While poverty may affect the worth of liberty, according to Rawls, poverty should not be seen as restricting liberty. Following the principles of justice, the basic structure will compensate for a lesser worth of liberty by maximising the primary goods available to the least advantaged. Rawls, 'The Basic Liberties and Their Priority', 41.

[1032] This also includes the right to hold (personal) property

of law[1033]."[1034] The basic liberties have a reconciling effect on people. Having equal access to them, people can regard one another "with mutual recognition on a footing of equality."[1035]

The basic liberties are presented as an outline to be further specified by the constitutional, legislative, and judicial stages.[1036] Rawls clarified how the basic liberties are impossible to specify independently of a given society's social, economic, and technological circumstances. Still, when we would expand our list of liberties to include liberties that are not critical, we risk weakening their position and protection.[1037] For this specification exercise, the two moral powers, and people's corresponding interest for self-determination and self-development, are the appropriate benchmark. The equal basic liberties exist so every citizen has equal access to the social conditions vital for their development and exercise. Rights and liberties gain significance when they are essentially involved in, or necessary to secure and protect, the adequate development and informed exercise of the moral powers.[1038] Whether the basic liberties have been adequately secured, must be assessed in light of those fundamental powers and not on the basis of the specific interests of people, or the success they have in achieving their own conception of the good.[1039]

The evaluation of algorithmic decision-making processes could be informed by the basic liberties in two ways. First, they may function as an additional trigger of caution for algorithmic decision-makers. Whenever an algorithmic system or structure touches upon the exercise of a basic liberty - whether it is through the outcomes produced, or due to the way they are imposed onto people – decision-makers could ask themselves whether their operations run the risk of limiting people's

---

[1033] This includes protection against arbitrary arrest and seizure. Rawls, *A Theory of Justice*, 1971, 61.

[1034] See for a specification: Rawls, 'The Basic Liberties and Their Priority', 5. See also: H. L. A. Hart, *Rawls on Liberty and Its Priority*, *Essays in Jurisprudence and Philosophy* (Oxford University Press), accessed 20 August 2020, http://oxford.universitypressscholarship.com/view/10.1093/acprof:oso/9780198253884.001.0001/acprof-9780198253884-chapter-11; Rawls, 'The Basic Liberties and Their Priority'; Rawls, *Justice as Fairness: A Restatement*.

[1035] Rawls, *Justice as Fairness: A Restatement*, 115.

[1036] Rawls, 'The Basic Liberties and Their Priority', 12.

[1037] Rawls, 10; John Rawls, *A Theory of Justice: Revised Edition* (Harvard University Press, 1999), 54, https://doi.org/10.2307/j.ctvkjb25m; Gabriel, 'Towards a Theory of Justice for Artificial Intelligence', 7.

[1038] For instance, with regard to people's capacity to develop a sense of justice, the equal political liberties and freedom of thought are necessary as they allow people to participate into the political landscape, effectively exercise their sense of justice and freely express their opinions doing so. In this regard, the fundamental rights enshrined within the European Convention of Human Rights and the Charter of Fundamental Rights of the European Union could also constitute basic liberties when they are truly essential towards the development of people's moral powers. Rawls, *Justice as Fairness: A Restatement*, 113. Rawls, 'The Basic Liberties and Their Priority', 7; Gabriel, 'Towards a Theory of Justice for Artificial Intelligence', 7–8.

[1039] Pogge, *John Rawls: His Life and Theory of Justice*, 87–90.

(equal) claim towards the enjoyment of the equal basic liberties. From a technical perspective, the often unavoidable presence of system bias could be particularly problematic in light of the requirement for people to have an equal right to the same scheme of basic liberties.[1040] In areas where public institutions make use of automated systems to govern and structure their decisions, system bias could be in direct conflict with the rule of law.[1041] Prejudiced predictive policing systems, which can moreover lead to arbitrary arrests, violate the demands of due process and equality before the law.[1042] Yet, looking beyond the public realm, and beyond the problem of biased (or otherwise prejudiced) decision-making systems, many more examples can be found of technologies that are (ab)used to directly interfere with people's equal basic liberties. The Cambridge Analytica scandal clearly illustrated how algorithms can be used to turn social media data into a political weapon.[1043] Special Rapporteur on freedom of religion and belief for the United Nations Ahmed Shaheed, warned against the effects technological advancements, including the use of pervasive inferential and emotion detection techniques, have on freedom of thought as they enable decision-makers to access and reveal a person's inner thoughts, and affect how people think, feel and behave.[1044] As aforementioned, for decision-makers to manipulate users in their exercise of choice through the deployment of such technologies could also be

---

[1040] See also: Gabriel, 'Towards a Theory of Justice for Artificial Intelligence'.

[1041] See for instance: Gentzel, 'Biased Face Recognition Technology Used by Government'.

[1042] See for instance: Paul Hayes, Ibo van de Poel, and Marc Steen, 'Algorithms and Values in Justice and Security', *AI & SOCIETY* 35, no. 3 (September 2020): 533–55, https://doi.org/10.1007/s00146-019-00932-9; Hildebrandt, 'Algorithmic Regulation and the Rule of Law'; Vogiatzoglou, 'Mass Surveillance,
Predictive Policing and the Implementation of the CJEU and ECtHR Requirement of Objectivity'; P. Jeffrey Brantingham, Matthew Valasik, and George O. Mohler, 'Does Predictive Policing Lead to Biased Arrests? Results From a Randomized Controlled Trial', *Statistics and Public Policy* 5, no. 1 (1 January 2018): 1–6, https://doi.org/10.1080/2330443X.2018.1438940; Andrew Guthrie Ferguson, 'Predictive Policing and Reasonable Suspicion', *EMORY LAW JOURNAL* 62 (n.d.): 68; Gentzel, 'Biased Face Recognition Technology Used by Government'.

[1043] Cadwalladr and Graham-Harrison, 'How Cambridge Analytica Turned Facebook "Likes" into a Lucrative Political Tool'. See for example: Huszár et al., 'Algorithmic Amplification of Politics on Twitter'; Dan Milmo and Dan Milmo Global technology editor, 'Twitter Admits Bias in Algorithm for Rightwing Politicians and News Outlets', *The Guardian*, 22 October 2021, sec. Technology, https://www.theguardian.com/technology/2021/oct/22/twitter-admits-bias-in-algorithm-for-rightwing-politicians-and-news-outlets; Engin Bozdag and Jeroen van den Hoven, 'Breaking the Filter Bubble: Democracy and Design', *Ethics and Information Technology* 17, no. 4 (2015): 249–65, https://doi.org/10.1007/s10676-015-9380-y; Seth Flaxman, Sharad Goel, and Justin M. Rao, 'Filter Bubbles, Echo Chambers, and Online News Consumption', *Public Opinion Quarterly* 80, no. S1 (2016): 298–320, https://doi.org/10.1093/poq/nfw006.

[1044] Ahmed Shaheed, The Freedom of Thought (5 October 2021); Available via: 'Freedom of Thought Increasingly Violated Worldwide, UN Expert Warns', OHCHR, accessed 30 September 2022, https://www.ohchr.org/en/press-releases/2021/10/freedom-thought-increasingly-violated-worldwide-un-expert-warns.

considered an impermissible act of domination.[1045] Likewise, algorithmic acts of oppression have the capacity to limit persons in their ability to associate, interact and engage with others. Finally, when decision-makers impose differentiation onto decision-subjects in ways that limit their basic liberties, or the equal enjoyment of these liberties vis-à-vis others, they must be able to provide a reasonable and objective justification for doing so. Given the essential nature of the basic liberties, any such intervention should be argued for in an accountable and transparent manner, and in ways that are subject to democratic review.[1046]

Second, Gabriel believes the basic liberties could be interpreted dynamically and adapted to the particular challenges brought forth as a result of modern-day societies' socio-technically mediated character.[1047] Following this evolutive approach, the two moral powers, and decision-subjects' interest towards self-determination and self-development, remain our benchmark. Yet, in light of the challenges posed by the algorithmic environment, the argument could be made that people should be given a general right to be free from algorithmic acts of oppression and domination as these acts are particularly inimical to the aforementioned interests.[1048]

Fair Equality of Opportunity

The first part of the second principle puts forward that social and economic inequalities should be attached to offices and positions, including economic and educational opportunities, and open to everyone under conditions of fair equality of opportunity.[1049]

[1045] Jeroen Van Den Hoven and Emma Rooksby, 'Distributive Justice and the Value of Information: A (Broadly) Rawlsian Approach', in *Information Technology and Moral Philosophy*, ed. Jeroen van den Hoven and John Weckert, Cambridge Studies in Philosophy and Public Policy (Cambridge: Cambridge University Press, 2008), 376–96, https://doi.org/10.1017/CBO9780511498725.019.

[1046] Jamie Grace and Roxanne Bamford, '"AI Theory of Justice": Using Rawlsian Approaches to Better Legislate on Machine Learning in Government', *SSRN Electronic Journal*, 2020, https://doi.org/10.2139/ssrn.3588256. See also: Van Den Hoven and Rooksby, 'Distributive Justice and the Value of Information'.

[1047] See also: Gabriel, 'Towards a Theory of Justice for Artificial Intelligence'. Rawls, 'The Basic Liberties and Their Priority', 6–7; Rawls, *Justice as Fairness: A Restatement*, 45.

[1048] See also: Chapter 6 "Socio-relational conceptualisations of equality"

[1049] Rawls ties fair equality of opportunity to pure procedural justice, which is obtained "when there is no independent criterion for the right result: instead, there is a correct or fair procedure such that the outcome is likewise correct or fair, whatever it is, provided that the procedure has been properly followed." Pure procedural justice can only exist "against the background of a just basic structure, including a just political constitution and a just arrangement of economic and social institutions." Fair equality of opportunity serves to insure pure procedural justice. Rawls, *A Theory of Justice*, 1971, 86–87. Wenar, 'John Rawls'. See also: Andrew Mason, 'Rawlsian Fair Equality of Opportunity', in *Levelling the Playing Field* (Oxford: Oxford University Press, 2006),

Rawls introduced the principle of fair equality of opportunity to "correct the defects of formal equality of opportunity – careers open to talents – in the system of natural liberty."[1050] Under the formal approach, people's access to opportunities, income and wealth are improperly influenced by morally arbitrary contingencies, i.e., their natural talents and abilities "as these have been developed or left unrealised, and their use favoured or disfavoured over time by social circumstances and such chance contingencies as accident and good fortune."[1051] This system runs the risk of concentrating property, wealth and political power in the hands of a select few.[1052] This risk can be mitigated through the introduction of the additional condition of fair chance: two people with the same level of talent and ability, and with the same willingness or motivation to use said talents and abilities should have an equal prospect of success, culture, and achievement, regardless of their initial social class of origin.[1053] Fair equality of opportunity thus tries to counteract the impact that differences in historical and social fortune have on life prospects, rather than the natural distribution of abilities talent.[1054]

The securitisation of fair equality of opportunity depends upon the combined efforts of all the political, legal, social and socio-technical institutions.[1055] For instance, in order to create fair equality of opportunity, regulators should not only prevent discrimination in the fields of education, but they should also ensure that the quality of education is roughly the same for everyone regardless of their income or class.[1056] How could fair equality of opportunity be used to guide algorithmic decision-making systems?

---

https://doi.org/10.1093/acprof:oso/9780199264414.003.0004; Larry A. Alexander, 'Fair Equality of Opportunity: John Rawls' Forgotten Principle', *Philosophy Research Archives* 11 (1985): 197–208, https://doi.org/pra19851111. Richard J. Arneson, 'Against Rawlsian Equality of Opportunity', *Philosophical Studies: An International Journal for Philosophy in the Analytic Tradition* 93, no. 1 (1999): 77–112; Paul Gomberg, 'Dilemmas of Rawlsian Opportunity', *Canadian Journal of Philosophy* 40, no. 1 (2010): 1–24.

[1050] Rawls, *Justice as Fairness: A Restatement*, 43.

[1051] Rawls, *A Theory of Justice*, 1971, 72.

[1052] Rawls, *Justice as Fairness: A Restatement*, 44.

[1053] Rawls, *A Theory of Justice*, 1971, 73; Rawls, *Justice as Fairness: A Restatement*, 43–44.

[1054] Rawls, *A Theory of Justice*, 1971, 72–73.

[1055] Rawls refers to fair equality of opportunity also as "liberal equality." Rawls, 73; Rawls, *Justice as Fairness: A Restatement*, 43–44. Rawls, *A Theory of Justice*, 1971, 73; Rawls, *Justice as Fairness: A Restatement*, 44. See also: Rawls on Fair Equality of Opportunity versus meritocracy: Rawls, *A Theory of Justice*, 1971, 106.

[1056] Wenar, 'John Rawls'. Alexander, 'Fair Equality of Opportunity', 198. Rawls, *Justice as Fairness: A Restatement*, 44.

First, even in the case algorithms do not directly govern people's opportunities to access certain positions, like a hiring algorithm would for instance, they have become increasingly integrated into, and actively co-shape, the major social institutions that govern people's access to advantage and disadvantage. Hence, regulators should ensure that algorithms operate in ways that are conducive in creating the right conditions under which fair equality of opportunity can be guaranteed and maintained. For example, algorithmic acts of oppression and domination undermine a person or group's social standing and make it more difficult for them to attain certain opportunities.[1057] Suppose a networking site for professionals only shows high-profile job advertisements to people that have graduated from elite schools. While the advertising algorithm might not necessarily interfere with the actual hiring practice of the company whose advertisements were shown – the company can still choose to promote talent and motivation during individual interviews - the advertising algorithm has nonetheless reduced the likelihood that people who come from a lower social class will actually apply for those positions in the first place.

Second, the principle of fair equality of opportunity may function as a direct guiding rule regarding the direct distribution of opportunities through algorithms, such as hiring or educational enrolment algorithms. On the one hand, these systems should be designed to offset the competitive advantage or disadvantage a person's social class of origin might have in the distribution of that particular opportunity. On the other hand, they should promote people's individual talents, abilities, motivations, and willingness.[1058] Enrolment algorithms that determine students' opportunity for higher education on the basis of general and historical grading patterns and previous school performance would be in violation with the principle of fair equality of opportunity.[1059] Such algorithms give insufficient account to students' individual achievements, potential and hard work. Furthermore, in countries where private schools are better funded and

---

[1057] The reverse is also true of course. As long as there is no correction of structural disadvantage within society, or no mechanism that truly facilitates fair equality of opportunity, algorithmic systems will tend to replicate the unfair division of opportunities within society. Hence, algorithmic fair equality of opportunity can only be achieved where fair equality of opportunity can be assured within society. In other words, the principle of fair equality of opportunity is also a warning sign towards algorithmic decision-makers: given the current circumstances, can algorithmic fair equality of opportunity be attained if our systems are not purposefully built to dismantle existing institutional barriers?

[1058] This conceptualisation of equality has also been used as a fair machine learning technique. Cynthia Dwork et al., 'Fairness Through Awareness', *ArXiv:1104.3913 [Cs]*, 19 April 2011, http://arxiv.org/abs/1104.3913; Matthew Joseph et al., 'Fair Algorithms for Infinite and Contextual Bandits', *ArXiv:1610.09559 [Cs]*, 29 October 2016, http://arxiv.org/abs/1610.09559.

[1059] Han, 'An Algorithm Shouldn't Decide a Student's Future'.

can be more selective, students attending these schools are placed at an advantage. In other words, the outcomes produced might also be classist.[1060]

The application of fair equality of opportunity onto the algorithmic environment is not always as straightforward. For one, social class of origin and class attitudes have a considerable influence on the way people discover, experience and develop their talents, motivations and preferences.[1061] If algorithmic decision-makers believe they judge people on the basis of their individual talents and abilities, they should also understand that those talents have been developed as a result of people's social environment.[1062] Where a person has had fewer opportunities to attain or develop a talent or skill due to their class of origin, their chance of success is limited against those who could. Proper fair equality of opportunity therefore requires decision-makers to take into account the socialisation of talents and abilities.[1063] Decision-makers should not be blind to social class of origin and historical disadvantage. They should map how the talents, motivations or abilities they have identified as a requirement for a particular position are shaped or hindered as a result of social circumstance.[1064]This exercise could help them account for, and overturn, the full impact of

---

[1060] See also: BBC News, 'A-Levels and GCSEs'.

[1061] In this sense, Rawls observes, the principle of fair opportunity can only be imperfectly carried out: "the extent to which natural capacities develop and reach fruition is affected by all kinds of social conditions and class attitudes." Rawls, *A Theory of Justice*, 1971, 74.

[1062] Arneson, 'Against Rawlsian Equality of Opportunity'; Lindblom, 'In Defense of Rawlsian Fair Equality of Opportunity', 242–43. See also: Fredman, 'Substantive Equality Revisited', 724; Bob Hepple, 'Discrimination and Equality of Opportunity—Northern Irish Lessons', *Oxford Journal of Legal Studies* 10, no. 3 (1990): 408–21, https://doi.org/10.1093/ojls/10.3.408.

[1063] Arneson, 'Against Rawlsian Equality of Opportunity'.

[1064] This could be contrasted to the technical approach that is "fairness through unawareness." Following such an approach, certain protected attributes, such as those associated with social class of origin, cannot be used within a decision-making system. Fairness through unawareness might be flawed as in certain cases the presence of such attributes could actually allow for the detection of unjustified practices, such as the identification of proxy discrimination on the basis of social class. Furthermore, those attributes might also be used to signal instances where social class could have had an effect on people's abilities. At the same time, it must be observed how the collection of protected or sensitive attributes presents greater privacy risks, and the collection of data on social class could lead to the oppressive act of over-exposure (See Chapter 6). Rather than argue in favour of the use or non-use of protected characteristics in technical systems, the condition advanced here would be a simpler one, i.e., for decision-makers to carefully consider how social class might impact the abilities they define as relevant for an opportunity or decision. See also discussions in: Pratik Gajane and Mykola Pechenizkiy, 'On Formalizing Fairness in Prediction with Machine Learning', *ArXiv:1710.03184 [Cs, Stat]*, 9 October 2017, http://arxiv.org/abs/1710.03184; Michael Veale and Reuben Binns, 'Fairer Machine Learning in the Real World: Mitigating Discrimination without Collecting Sensitive Data', *Big Data & Society* 4, no. 2 (December 2017): 205395171774353, https://doi.org/10.1177/2053951717743530.8 Berendt and Preibusch, 'Better Decision Support through Exploratory Discrimination-Aware Data Mining'.

social class has in their distribution of opportunities, rather than have it unjustifiably result in a loss thereof.[1065] In this sense, the requirement of fair equality of opportunity would issue a call towards reflection aimed to overcome, rather than maintain, social hierarchies.[1066]

Decision-makers should thus carefully reflect upon those talents, abilities and motivations actually matter for a given position or opportunity, and how those requirements can be shaped and influenced by social class. At the same time, they should also exclude as differentiation grounds, those attributes and traits that are not reflective of talent, ambition, or willingness. Once the relevant criteria have been defined and formalised however, decision-makers should continue to monitor whether *in practice* the chosen requirements do not lead to the exclusion of people based on their social class.

The principle of fair equality of opportunity may also guide the use of novel and non-representational differentiation grounds. Still, it would be difficult to evaluate the actual validity of those grounds as being reflective of a person's ambitions or talents.[1067] Imagine a recruitment agency has built an online platform so all job seekers can fill in their resumes in a uniform format. Having analysed the data obtained from this registration process, the agency discovers there is a relationship between a job seeker's keystroke patterns and their client's level of satisfaction. Though at first sight, keystroke patterns hold no direct relationship with a person's capacity to perform a job, the recruitment agency argues otherwise. Knowledge discovery techniques have uncovered this attribute as a reliable indicator, and it will be used to decide who to send to high-end clients. As seen in the previous Chapter, such conundrums may be resolved by allowing decision-subjects to contest decisions and have their individual circumstances be taken into consideration.[1068] Alternatively, Wachter has argued that there must be at least an intuitive link

---

[1065] See also: Lindblom, 'In Defense of Rawlsian Fair Equality of Opportunity'. Difficulties may again arise. For example, is affirmative action allowed under fair equality of opportunity? Thomas Nagel, 'John Rawls and Affirmative Action', *The Journal of Blacks in Higher Education*, no. 39 (2003): 82, https://doi.org/10.2307/3134387.

[1066] Lindblom, 'In Defense of Rawlsian Fair Equality of Opportunity', 262.

[1067] According to Lee et al., the principle of fair equal opportunity allows algorithmic decision-makers to classify and differentiate between people on the basis of their natural characteristics, talents, and their preferences, but it is opposed against the use of certain socio-economic markers as these would indicate social class of origin. Under this view, differentiation grounds generated through behavioural analytics indicative of people's preferences could be used as a valid differentiation ground. Michelle Seng Ah Lee, Luciano Floridi, and Jatinder Singh, 'Formalising Trade-Offs beyond Algorithmic Fairness: Lessons from Ethical Philosophy and Welfare Economics', *AI and Ethics* 1, no. 4 (November 2021): 529–44, https://doi.org/10.1007/s43681-021-00067-y.

[1068] See also Chapter 6, Section 6.2.2 The Problematic Nature of Generalisations"

between the criteria used and the decision at hand.[1069] An interesting solution has been offered by Lindblom. He believes that what matters most within the framework of fair equality of opportunity is to evaluate how a given differentiation ground influences someone's social standing: "The core aim of [fair equality of opportunity] is to secure the equal standing of citizens, and its measure of unjust discrimination is this ideal."[1070] Under this view, and rather than having to verify the specific nature of the differentiation ground and its relationship to individual talent and ability, fair equality of opportunity mandates decision-makers to consider the negative impact their decision-making methods could have on a person's social standing. The latter would require them to make a contextual analysis and think about their social influence, power, and impact on people's mutual relationships, i.e., the risk of socio-relational egalitarian harms.

The Difference Principle

The principle of fair equality of opportunity constitutes a first step in correcting a system of natural liberty but is imperfect on its own. While it eliminates "the influence of social contingencies, it still permits the distribution of wealth and income to be determined by the natural distribution of abilities and talents."[1071] As Rawls observes: "there is no more reason to permit the distribution of income and wealth to be settled by the distribution of natural assets than by historical and social fortune."[1072]

THE DIFFERENCE PRINCIPLE - In order to mitigate the impact of natural assets, as well as fortune, Rawls invokes the "difference principle": in order to be justified all social and economic inequalities (inequalities in wealth and income, as well as inequalities in the social bases of self-respect) should be to greatest benefit of those that are the least-advantaged within society.[1073] Even though people can make use of their talents and abilities to generate greater income and wealth, the greater economic reward they produce must be to the benefit of those least-advantaged and maximise the latter's level of primary goods. [1074] The difference principle is

[1069] Wachter, 'The Theory of Artificial Immutability', 40.

[1070] Lindblom, 'In Defense of Rawlsian Fair Equality of Opportunity', 240.

[1071] Rawls, *A Theory of Justice*, 1971, 74.

[1072] Rawls, 74.

[1073] One could wonder however, why the principle of fair equality of opportunity should be given priority over the difference principle given that both inequalities following social class and natural endowments are morally arbitrary. On this point, see also: Mason, 'Rawlsian Fair Equality of Opportunity'. Alexander, 'Fair Equality of Opportunity'.

[1074] Lamont and Favor, 'Distributive Justice'. While our social class of origin, native endowments and good or bad fortune may be morally arbitrary, the difference principle does not say that people cannot deserve the positions they hold in

positively aimed towards an ideal of deep social unity: "In a society that satisfies the difference principle, citizens know that their economy works to everyone's benefit, and that those who were lucky enough to be born with greater natural potential are not getting richer at the expense of those who were less fortunate."[1075]

THE LEAST ADVANTAGED – The least advantaged are likely to be found at the intersections, among those least favoured by the three broad contingencies of life: "persons whose family and class origins are more disadvantaged than others, whose natural endowments have permitted them to fare less well, and whose fortune and luck have been relatively less favourable."[1076] In themselves these contingencies do not define the least advantaged, nor can the least-advantaged be identified by reference to specific salient traits or natural features, such as ethnicity or gender, alone.[1077] The least fortunate are defined in relation to the primary goods.[1078] Because the basic liberties and fair equality of opportunity should have already been guaranteed to all, Rawls argues that the least advantaged are likely the persons who belong to the "income class[es] with the lowest expectations."[1079] While the least advantaged cannot be defined solely in relation to their income and wealth, they can also not be identified independent of their income and wealth.[1080]

Decision-makers must compare different policy options and select those options that make the

---

life, or that they do not deserve the skills and abilities they have realised throughout their lifetime. Rawls, *Justice as Fairness: A Restatement*, 78.

[1075] Leif Wenar, 'John Rawls', in *The Stanford Encyclopedia of Philosophy*, ed. Edward N. Zalta, Spring 2017 (Metaphysics Research Lab, Stanford University, 2017), https://plato.stanford.edu/archives/spr2017/entries/rawls/. On an economic level, the difference principle has been inspired by the world view that wealth is most readily increased where productivity is rewarded by greater income. On a normative level, the difference principles is motivated by the belief that people are (moral) equals, regardless of the contingencies that affect their life's prospects. Lamont and Favor, 'Distributive Justice'.

[1076] Rawls, 'A Kantian Conception of Equality'.

[1077] As Rawls observes: "Even supposing, for example, that it turns out, as common-sense political sociology might suggest, that the least advantaged, identified by income and wealth, include many individuals born into the least-favoured social class of origin, and many of the least (naturally) endowed and many who experience more bad luck and misfortune […], nevertheless those attributes do not define the least advantaged. Rather, it happens that there may be a tendency for such features to characterize many who belong to that group." Rawls, *Justice as Fairness: A Restatement*, 59, footnote 26.

[1078] Rawls, 'A Kantian Conception of Equality'.

[1079] Rawls, *Justice as Fairness: A Restatement*, 59. Rawls does specify that within his theoretical conception of justice every person has a "normal range" of physical and mental capacities. Rawls, 64–65.

[1080] Rawls, *Justice as Fairness: A Restatement*, at footnote 26.

least-advantaged best off.[1081] The Difference Principle "applies in the first instance to the main public policies and principles that regulate social and economic inequalities." While the Difference Principle was not intended to apply to particular or local transactions or small-scale decisions taken within a system, it could nonetheless inform the regulation algorithmic decision-making systems.[1082] First, algorithms have become an integral part of the social world and the institutions that structure and govern people's everyday lives.[1083] Algorithmic structures co-shape social and economic inequality and the outcomes they produce should be to the benefit of those least-advantaged.[1084] Yet, I believe the same could be said for individualised or isolated decision-making instances. Algorithms function on the basis of higher-level instructions that capture a totality of decision-subjects. These instructions function as a type of policy that will determine how, within a given context, algorithms will distribute wealth, resources, and social standing among decision-subjects. Following this interpretation, and given the availability of various decision models and designs, decision-makers should make a comparison between these alternatives in light of the inequalities they produce and opt to build those systems that make the least-advantaged most well off.[1085] In this context, the Difference Principle could take on a more transformative function, towards the building of systems that are able to effectuate change.

The identification of the "least advantaged" within a given society is a challenging task. The Dutch Child-Benefit scandal might offer a clear example of an algorithmic system that failed to respect the difference principle. Dutch tax authorities falsely accused thousands of parents and caregivers of childcare benefit fraud. The tax authorities made use of a black-box risk classification model that included self-learning algorithms. Amnesty International observed how "different layers of inequality were created through discrimination based on race and ethnicity in the use of the risk

[1081] Rawls, 59–60. By maximising a person's access to wealth and income, it is believed their enjoyment of other primary goods can increase too. For instance, it will allow the least-advantaged too better enjoy their basic liberties. It is this focus on income and wealth, that has led Rawls being criticised among others by capability theorists. For instance, a member of an oppressed group might have accumulated great wealth, but still find themselves unable to meaningfully participate in society. For a brief overview of arguments against the Difference Principle, reference can be made to: Lamont and Favor, 'Distributive Justice'.

[1082] Rawls, 'A Kantian Conception of Equality'.

[1083] David Beer, 'The Social Power of Algorithms', *Information, Communication & Society* 20, no. 1 (2017): 1–13, https://doi.org/10.1080/1369118X.2016.1216147.

[1084] See also Section below: Social Bases of Self-Respect

[1085] See also: Gabriel, 'Towards a Theory of Justice for Artificial Intelligence', 9–10.

classification model, combined with the policies and practices that more often affected people with a lower economic status, who often belonged to an ethnic minority."[1086]

In the definition of the least-advantaged, Rawls emphasises the importance of wealth and income. To define the least advantaged through the lens of those primary goods alone, however, might limit the evaluative potential of the difference principle. For example, one might fail to protect (oppressed) social groups who face hardship and inequality on a socio-relational level because they are (more or less) well-off economically. Hence, the impact algorithmic decisions have on people's social standing (represented through the primary good of the social bases of self-respect) could be further accentuated.[1087]

In addition, the notion of the least-advantaged might need to be approached with a certain degree of flexibility in order to mitigate the various socio-relational and socio-economic harms different systems can impose onto diverse groups. Take again the Austrian AMS employment algorithm. As above-mentioned, women, disabled people, and people over thirty were disfavoured by the system and were less likely to obtain direct unemployment support.[1088] Persons who find themselves at the intersection of the aforementioned identities might belong to the least-advantaged group on the job market. But are they still the least-advantaged group outside of this context and hence deserving of protection under the Difference Principle?[1089]

(The Social Bases of) Self-Respect

Among the primary goods subject to distribution are the social bases of self-respect. They are those aspects of the political, economic, and social institutions that allow people to develop a sense

---

[1086] 'Xenophobic Machines: Discrimination through Unregulated Use of Algorithms in the Dutch Childcare Benefits Scandal.' See also: Huisman, 'SyRI, het fraudesysteem van de overheid, faalt'; van Bekkum and Borgesius, 'Digital Welfare Fraud Detection and the Dutch SyRI Judgment'; Vervloesem, 'How Dutch Activists Got an Invasive Fraud Detection Algorithm Banned'; Jan Kleinnijenhuis, 'Hoe de Belastingdienst lage inkomens profileerde in de jacht op fraude', Trouw, 22 November 2021, https://www.trouw.nl/gs-bbb66add; 'Xenophobic Machines: Discrimination through Unregulated Use of Algorithms in the Dutch Childcare Benefits Scandal.'

[1087] See also Section below "The Social Bases of Self-Respect)

[1088] Wagner et al., 'Der AMS-Algorithmus.'; Lopez, 'Bias Does Not Equal Bias'; Lopez, 'Reinforcing Intersectional Inequality via the AMS Algorithm in Austria'; Kayser-Bril, 'Austria's Employment Agency Rolls out Discriminatory Algorithm, Sees No Problem'. Allhutter et al., 'Algorithmic Profiling of Job Seekers in Austria'.

[1089] Of course, there are limits to the notion of the least-advantaged. Most decision-making systems allocate advantages and disadvantages among people. It is not because a given group is allocated a particular disadvantage (e.g., non-ideal applicants in a hiring decision) that they can claim a violation of the difference principle.

of self-worth and confidence during the development and exercise of their moral powers and the pursuit of their goals.[1090] Self-respect, according to Rawls, comprises two key components. First, self-respect refers to a person's sense of value and their conviction that their idea of the good life is worthwhile pursuing.[1091] Second, self-respect instils within a person the self-confidence that they can fulfil their life goals.[1092]

Having become an intrinsic element of the basic institutions, algorithmic structures and decisions should nurture, rather than impede person's self-respect. As illustrated in the previous Chapter, algorithmic forms of oppression and domination actively erode a decision-subject's self-worth and confidence through actions that undermine a person or group's social standing and socio-economic prospects, their capacity towards self-development and self-determination.[1093]

Drawing on the work of Manders-Huits, Hoffman believes algorithmic categorisation and classification acts represent several identity-related risks that thwart a person's self-respect.[1094] First, classifications represent a false confidence that everything can be expressed in or through data, which leads to a simplified view of reality and the person, and denies those captured by the classification their self-authorship.[1095] Second, classifications within the digital environment can be persistent and tied to one's person over long periods of time. A classification might have been wrong, or the information caught by the classification might no longer correspond to a person's actual circumstances or convictions. Yet, once generated, (false) information can endure, be easily disseminated, and difficult to remove.[1096] A person might have been misidentified as a security threat. Yet because this label persists somewhere in a database, they now struggle to find a steady job.[1097] Finally, classifications are not purely descriptive, they can be co-constitutive: people might

---

[1090] Rawls, *Justice as Fairness: A Restatement*, 60.

[1091] Rawls, *A Theory of Justice*, 1971, 440.

[1092] Rawls, 440.

[1093] See Chapter 6 "Socio-Relational Conceptualisations of Equality"

[1094] Anna Lauren Hoffmann, 'Rawls, Information Technology, and the Sociotechnical Bases of Self-Respect', The Oxford Handbook of Philosophy of Technology, 10 May 2022, https://doi.org/10.1093/oxfordhb/9780190851187.013.15; Noëmi Manders-Huits, 'Practical Versus Moral Identities in Identity Management', *Ethics and Information Technology* 12, no. 1 (2010): 43–55, https://doi.org/10.1007/s10676-010-9216-8. See also Chapter 6, Section 6.2.2 "The Problematic Nature of Generalisations."

[1095] This is referred to as computational reductionism. Manders-Huits, 'Practical Versus Moral Identities in Identity Management', 51–52.

[1096] Manders-Huits, 52; Hoffmann, 'Rawls, Information Technology, and the Sociotechnical Bases of Self-Respect', 243.

[1097] Bruno Struys, 'Plots stond Bilal op een lijst van de veiligheidsdiensten: "Ik werd bang van mezelf toen ik het rapport las"', De Morgen, 29 September 2022, https://www.demorgen.be/gs-b11dfd65.

"implicitly shape their (moral) identities according to the possibilities and constraints offered by available labels." Suppose a system believes a decision-subject belongs to a particular demographic. The options presented to the decision-subject are now changed in order to correspond to this newly imposed identity. Being constrained in their actions and behaviours, the decision-subject now conforms to, reinforces, and behaves in accordance with, the label's expectations.[1098] Algorithmic classifications might also inhibit people upon their discovery. Imagine it is disclosed to a person that they were classified as running an elevated risk of cancer. Such information can affect a person's in the construction of their identity in ways that undermine their sense of worth and self-confidence, as well as alter the concrete life plans they wanted to pursue.[1099] The choices embedded within classification acts can therefore hinder or promote people in their development of self-respect.[1100] They too could constitute an algorithmic act of oppression or domination, and hence something algorithmic decision-makers should be mindful of.

[1098] This is referred to as dynamic nominalism. In itself a notion taken from Ian Hacking. Noëmi Manders-Huits, 'Practical versus Moral Identities in Identity Management', *Ethics and Information Technology* 12, no. 1 (March 2010): 52–53, https://doi.org/10.1007/s10676-010-9216-8; Ian Hacking, 'Making up People', in *Reconstructing Individualism: Autonomy, Individuality, and the Self in Western Thought.*, ed. T.C. Heller, M. Sosna, and D.E. Wellbery (Stanford: Stanford University Press, 1986), 222–36. Lawrence Lessig, *Code*, Version 2.0 (New York: Basic Books, 2006), 154.

[1099] Manders-Huits, 'Practical versus Moral Identities in Identity Management', 53.

[1100] Hoffmann, 'Rawls, Information Technology, and the Sociotechnical Bases of Self-Respect', 245.

7.3 Equality as Capability

Primary goods, and wealth and income in particular, are not necessarily good indicators for the kind of life someone can or wants to lead or achieve.[1101] Resource-based approaches tend to overlook the critical differences that exist between individuals in their capacity to convert those resources into the realisation of life goals.[1102] As an alternative metric, Amartya Sen and Martha Nussbaum proposed a different currency for justice: equality of capability.

If a society wants to compare persons and groups in terms of their advantage and privilege, and evaluate to what extent differences between people - whether they be in status or material - are justified, it should measure people in terms of their overall capabilities.[1103] Capabilities constitute the substantive freedoms or opportunities people have in order to achieve alternative beings and doings (functionings).[1104] Sen observes that the capability approach "can be seen as a natural extension of Rawls' concern with primary goods, shifting attention from goods, to what goods *do* to human beings."[1105] This shift in focus might not be radical, but it is a shift nonetheless.[1106] For instance, the capability approach will pay increased attention to the diversity of human needs and

---

[1101] With regard to disabilities, Nussbaum takes an even stronger stance. She argues that the social contract tradition as such, which does also include Rawls' original position, is not capable of showing sufficient sensitivity towards the value and moral worth of people with disabilities. See among others: Nussbaum, *Frontiers of Justice: Disability, Nationality, Species Membership*, 2006; Martha Nussbaum, 'The Capabilities of People with Cognitive Disabilities', *Metaphilosophy* 40, no. 3/4 (2009): 331–51; Michael Bérubé, 'Equality, Freedom, and/or Justice for All: A Response to Martha Nussbaum', *Metaphilosophy* 40, no. 3–4 (2009): 352–65, https://doi.org/10.1111/j.1467-9973.2009.01595.x.

[1102] Martha Nussbaum, 'Women and Equality: The Capabilities Approach Special Issue: Women, Gender and Work (Part I)', *International Labour Review* 138, no. 3 (1999): 232. This applies to Rawls, but also Dworkin's equality of resources.

[1103] Sen, *The Idea of Justice*, 253.

[1104] Robeyns, 'The Capability Approach'. Sen, *The Idea of Justice*, 253. For an analysis capability approach in light of the history of philosophy, see also: Martha C. Nussbaum, 'The Capabilities Approach and the History of Philosophy', in *The Cambridge Handbook of the Capability Approach*, ed. Enrica Chiappero-Martinetti, Siddiqur Osmani, and Mozaffar Qizilbash, 1st ed. (Cambridge University Press, 2020), 13–39, https://doi.org/10.1017/9781316335741.003.

[1105] For instance, Rawls highlights the importance of income, rather than what income does. Sen, 'Equality of What?', 218–19. Whereas the Capability Approach has been considered as a complement to Rawls, Nussbaum points out that it was above all meant as an alternative to the classic economic-Utilitarian paradigm. Nussbaum, *Frontiers of Justice: Disability, Nationality, Species Membership*, 2006, 70.

[1106] Robeyns, 'The Capability Approach'. More information on how the capability approach opposes itself against other theories of justice can be found within the work of Nussbaum. Whereas *Women and Human Development* can be seen as a critique against utilitarianism, and *Frontiers of Justice* against the deficits of social contract theories. Martha C. Nussbaum, *Women and Human Development: The Capabilities Approach*, 12th ed. (New York: Cambridge University Press, 2000); Nussbaum, *Frontiers of Justice: Disability, Nationality, Species Membership*, 2006.

the personal and contextual circumstances that shape people's lives: "By focusing on what people are able to do and be, rather than merely on the distribution of goods and resources, the capability approach recognizes the diversity of people's ability to convert those resources and goods into real opportunities and achievements - the kind of life they are effectively able to lead."[1107]

Justice should be concerned with those types of activities people can do and what kind of person people can be.[1108] Taking the latter as core goal of justice, Martha Nussbaum adapted and developed the capability approach into a (partial) theory of justice in an effort to "provide the philosophical underpinning for an account of core human entitlements that should be respected and implemented by the governments of all nations, as a bare minimum of what respect for human dignity requires."[1109] Her theory has the following key elements.[1110] First, people must be treated as an end, and in doing so, society should be concerned with the real opportunities people have. Second, a capability-oriented theory of justice should be focused on choice or freedom: provided that people are given substantial opportunities and freedoms, people should be given respect in their power of self-definition and determination. Third, justice is pluralist (and liberal) about value. There is no single conception of the good and people ought to be given the freedom necessary to pursue their own good. Finally, justice should be concerned with "entrenched social injustice and inequality" and especially with those "capability failures that are the result of discrimination or marginalisation."[1111] Governments and public policy are tasked to improve "the quality of life for all people, as defined by their capabilities."[1112] This approach culminates into a

---

[1107] These contextual circumstances can be a person's abilities and the political, social, economic and socio-technical environment in which a person finds themselves. Robeyns, 'The Capability Approach'. See also: Amartya Sen, 'Essays and Reviews', in *1959-2002* (Princeton University Press, 2014), chap. Equality Re-examined, https://doi.org/10.1515/9781400848393-065. Nussbaum, *Frontiers of Justice: Disability, Nationality, Species Membership*, 2006, 88.

[1108] Robeyns, 'The Capability Approach'.

[1109] Nussbaum, *Frontiers of Justice: Disability, Nationality, Species Membership*, 2006, 70.

[1110] Nussbaum, *Creating Capabilities: The Human Development Approach*, 18–19.For a rival description ,see also: CA Ingrid Robeyns, 'Capabilitarianism', *Journal of Human Development and Capabilities* 17, no. 3 (2 July 2016): 397–414, https://doi.org/10.1080/19452829.2016.1145631. Robeyns considers that only Nussbaum's consideration that the CA should focus on people as ends, and that the CA is pluralist in value are essential to the CA. Robeyns distinguishes 12 characteristics to the CA, 6 of which can be considered "core" to every CA (including Nussbaum's): 1) capabilities and functionings as core concepts; 2) the means-end distinction; 3) sensitivity towards differences in people's conversion functions, which allow them to transform resources into functionings; 4) attention to human diversity (and differing needs); 5) structural constraints, such as laws and social norms, as limiting people's capabilities; 6) the acknowledgement of agency.

[1111] Nussbaum, *Creating Capabilities: The Human Development Approach*, 18–19.

[1112] Nussbaum, 19.

list of central human capabilities, which are "all implicit in the idea of a life worthy of human dignity."[1113] Firmly rooted within the (political) liberal tradition, the central human capabilities are seen as "a source of political principles for a liberal pluralistic society."

Drawing from Nussbaum's capability theory, in the following Sections, I evaluate algorithmic inequalities from a capability perspective. In the first sub-section, I highlight the conceptual difference between capabilities and functionings. In the second section, I explain how the capability theory is tied to the notion of human dignity. Finally, the central human capabilities are put forward as a benchmark through which we could evaluate and regulate. algorithmic decision-making practices.

[1113] Nussbaum, *Frontiers of Justice: Disability, Nationality, Species Membership*, 2006, 70. Nussbaum, 70. The list can be placed within a political liberal tradition in the sense that, similar to Rawls, it does not propose a specific definition of the good. On this topic, Nussbaum, in reference to her early work on Aristotle, observed: "*it would be misleading to think that I see my own theory as a theory of good human functioning closely analogous to a theory of moral virtue. Such a reading would ignore the fact that, in my view, the appropriate political goal is not functioning, but simply capability. And it would ignore the fact that, since 1994 at any rate, I hold a form of political liberalism that makes it inappropriate for any particular comprehensive conception of ethical value to be endorsed by politics. Aristotle sees the production of virtuous functioning as among the legitimate ends of politics; I do not."* Martha C. Nussbaum, 'Aristotle, Politics, and Human Capabilities: A Response to Antony, Arneson, Charlesworth, and Mulgan', *Ethics* 111, no. 1 (2000): 124, https://doi.org/10.1086/233421. Other scholars have argued however that a return to the Aristotelian roots may be more appropriate for the CA. Severine Deneulin, 'Recovering Nussbaum's Aristotelian Roots', *International Journal of Social Economics* 40, no. 7 (2013): 624–32, https://doi.org/10.1108/IJSE-2012-0127. Rutger Claassen, 'Human Dignity in the Capability Approach', in *The Cambridge Handbook of Human Dignity*, ed. Marcus Duwell et al. (Cambridge: Cambridge University Press, 2014), 240–49, https://doi.org/10.1017/CBO9780511979033.030. John M. Alexander, 'Social Justice and Nussbaum's Concept of the Person*', in *Capabilities, Gender, Equality: Towards Fundamental Entitlements*, ed. Flavio Comim and Martha C. Nussbaum (Cambridge: Cambridge University Press, 2014), 414–36, https://doi.org/10.1017/CBO9781139059138.021.

7.3.1 Capabilities and Functionings[1114]

Nussbaum makes a distinction between three types of capabilities: basic (or innate), internal and combined capabilities.[1115]

The basic capabilities refer to the innate capacities that a person has and are necessary to develop and train more advanced capabilities.[1116] They are "innate powers that are either nurtured or not nurtured."[1117] For instance, a new-born child might have the (rudimentary) capability for speech, love, practical reason, work, etc.[1118] Internal capabilities are trained and developed traits and abilities. Though most people have an innate human capacity to talk, they must train to turn this basic capability into the more advanced capability of singing. The internal capabilities are cultivated in interaction with a person's external environment, such as their social, economic, and political surroundings, and governments should promote their development.[1119] The combined capabilities take into account the entire context that is relevant to a person: they are the totality of opportunities and freedoms a person has to choose and act given their specific political, economic and social environment.[1120] They are "a combination of a person's internal capabilities with the suitable external conditions for the exercise of the function."[1121] Whereas (basic and) internal capabilities may be a starting point for people's enjoyment of their freedoms and opportunities, the combined capabilities constitute their real opportunities or substantive freedoms they have.[1122]

[1114] The language used by Nussbaum is distinct from the vocabulary used within the CA disciplines. In most CA literature, "capabilities aim to capture what people are *able* to do and to be" (opportunities), whereas the notion functionings "point[s] at the corresponding achievements" (outcomes). Capabilities correspond to real freedoms and opportunities. Nussbaum however reserves the term "combined capabilities" for the latter. Robeyns, 'Capabilitarianism', 405–6. Ingrid Robeyns, *Wellbeing, Freedom and Social Justice: The Capability Approach Re-Examined* (Open Book Publishers, 2017), 39–40, https://doi.org/10.11647/OBP.0130.

[1115] For an early account of this distinction, see also: Martha Nussbaum, 'Nature, Function, and Capability: Aristotle on Political Distribution', *Oxford Studies in Ancient Philosophy: Supplementary Volume*, 1 January 1988, https://chicagounbound.uchicago.edu/book_chapters/782.

[1116] Nussbaum, *Creating Capabilities: The Human Development Approach*, 23. Nussbaum, *Women and Human Development: The Capabilities Approach*, 84–85.

[1117] Nussbaum, *Creating Capabilities: The Human Development Approach*, 23.

[1118] Nussbaum, *Women and Human Development: The Capabilities Approach*, 84–85.

[1119] Nussbaum, *Creating Capabilities: The Human Development Approach*, 21.

[1120] Nussbaum, 21.

[1121] Nussbaum, *Women and Human Development: The Capabilities Approach*, 84–85; Nussbaum, *Creating Capabilities: The Human Development Approach*, 20–21.

[1122] Robeyns, *Wellbeing, Freedom and Social Justice*, 93–94.

Functionings are the "active realization of one or more capabilities."[1123] Capabilities exist so they can be translated into functionings: "the notion of functionings gives the notion of capability its end-point."[1124] For example, while a person's being educated or wanting to travel may be the desired functioning, their ability to choose education or travel in a truly free manner, would be the corresponding combined capability.[1125] The concept of combined capability captures the following: could person A do X, if they wanted to do X.[1126]

Capabilities express the freedom to choose or the opportunity to select functionings, and "to promote capabilities is to promote areas of choice."[1127] They also reflect the importance of people's right to self-determination and self-development. On the one hand, people should have freedom to pursue their own ends in accordance with their choice (self-determination). On the other hand, a just society provides the suitable external conditions so that people can transform and exercise their skills in socially meaningful and recognised ways (self-development).[1128] Following Young, the conditions that constrain people in their capabilities can also be tied to the (institutional) organisation of power, status, and communication.[1129] A just society should therefore not only ensure that people can transform their innate capabilities into internal capabilities but also remove the external obstacles that prevent people to function in accordance with their capabilities. A society may do well at producing the former, it can fail to do the latter. A student may indeed have the required internal capabilities to successfully graduate university, but these capabilities cannot flourish if educational institutions employ enrolment algorithms that prioritise historical school records, rather than students' actual learning capacities.[1130] Likewise,

---

[1123] Nussbaum, *Creating Capabilities: The Human Development Approach*, 24–25.

[1124] Nussbaum, 25.

[1125] In this context, Robeyns also makes a distinction between universal functionings that should be guaranteed to all and context-dependent functionings that are mainly dependent on the existing social structures. For instance, Robeyns notes that while owning a house is a context-dependent functioning, "being sheltered from the elements" should arguably be achieved for every living person regardless of their social circumstance. To some extent then, to safeguard for every person the essence or core of the fundamental human rights could be seen as "universal" functionings.

[1126] Robeyns, *Wellbeing, Freedom and Social Justice*, 39–40. Robeyns sums up the difference thusly: "The distinction between functionings and [combined] capabilities is between the realized (achievements) and the effectively possible (freedoms and opportunities to choose from)." Robeyns, 39. Robeyns, 91. In practice however the use of both notions is not always as clear. For more information on both concepts, reference can be made to Robeyns, chaps 3, Clarification.

[1127] Nussbaum, *Creating Capabilities: The Human Development Approach*, 25.

[1128] In this regard, Young herself has observed how "equality as capabilities" is most closely associated to her understanding of self-development. Young, *Inclusion and Democracy*, 31–32.

[1129] Young, 31–32.

[1130] See also Chapter 2 Section "The Evaluative Advantage of Equality"; Chapter 6 Section 6.1.5 "Algorithmic Oppression"

transgender or gender nonconforming people have the required internal capabilities for travelling, but they face numerous obstacles in doing so: identification documents may not represent their gender expression or body scanners at the airport may not recognise the diversity of the human body when built on gender-stereotyping attributes, etc.[1131]

7.3.2 Equal Human Dignity: Equality of Capability

Equality of capability is motivated by a deep respect for equal human dignity.[1132] According to Nussbaum, "the basic intuitive idea [underlying a capability-oriented theory of justice is to] begin with a conception of the dignity of a human being, and of a life that is worthy of that dignity – a life that has available in it 'truly human functioning.'"[1133]

[1131] " National Center for Transgender Equality, 'Airport Security', *National Center for Transgender Equality*, accessed 29 April 2022, https://transequality.org/know-your-rights/airport-security; Medina, 'TSA's Body Scanners Are Gender Binary. Humans Are Not.'; Story by Allison Hope Hayes Illustrations by India, 'Traveling as a Trans Person: It's Complicated', CNN, accessed 29 April 2022, https://www.cnn.com/travel/article/transgender-travel-complications-tips/index.html.

[1132] Rawls too linked the principles of justice to dignity: "once the conception of justice is on hand, however, the ideas of respect and of human dignity can be given more definite meaning." Rawls, *A Theory of Justice*, 1971, 586. Beyond dignity Nussbaum's writings on compassion and moral sentiments could be of value too. However, this specific topic, I shall not discuss here. Martha C. Nussbaum, 'Human Functioning and Social Justice: In Defense of Aristotelian Essentialism', *Political Theory* 20, no. 2 (1992): 202–46; Martha C. Nussbaum, ed., 'Compassion: The Philosophical Debate', in *Upheavals of Thought: The Intelligence of Emotions* (Cambridge: Cambridge University Press, 2001), 354–400, https://doi.org/10.1017/CBO9780511840715.009; Martha C. Nussbaum, ed., 'Compassion and Public Life', in *Upheavals of Thought: The Intelligence of Emotions* (Cambridge: Cambridge University Press, 2001), 401–54, https://doi.org/10.1017/CBO9780511840715.010. Alexander, 'Social Justice and Nussbaum's Concept of the Person*'. John M. Alexander, *Capabilities and Social Justice: The Political Philosophy of Amartya Sen and Martha Nussbaum* (Ashgate, 2008), 141.

[1133] The central capabilities exert a moral claim. The list of central human capabilities is informed by Marxist ideal that sees a human being "in need of a totality of human-life activities." Nussbaum, *Frontiers of Justice: Disability, Nationality, Species Membership*, 2006, 74. Nussbaum's understanding of dignity draws from the philosophy of Aristotle and Marx, as well as Kant and the Stoics. For an exploration of dignity in relation to other philosophical strands, see: Martha Nussbaum, 'Human Dignity and Political Entitlements', *Human Dignity and Bioethics: Essays Commissioned by the President's Council on Bioethics*, 1 January 2008, https://chicagounbound.uchicago.edu/book_chapters/202; Nussbaum, 'The Capabilities Approach and the History of Philosophy'. Nussbaum, *Women and Human Development: The Capabilities Approach*, 74. For a more historical overview, see also: Nussbaum, 'The Capabilities Approach and the History of Philosophy'; Martha C. Nussbaum, *The Cosmopolitan Tradition: A Noble But Flawed Ideal* (Cambridge, Massachusetts ; London, England: Belknap Press: Harvard University Press, 2019).

The concept of dignity is given specific shape through its connection with other notions.[1134] First, dignity is linked to agency. Dignity dictates that policy choices should promote and protect agency and choice, "rather than infantilise people and treat them as passive recipients of benefit." Second, dignity is linked to respect. Beings with dignity demand respect.[1135] When people have respect for another person's dignity, this should prevent them from using others as a tool or means. Respect for human dignity demands of a society to put in place a set of substantial freedoms in which people can develop and exercise their capacities and choice, which allows people to view themselves (and others) as an end. Finally, due to their shared human dignity, people have an equal entitlement to equality of capability. This is an essential social goal and equality of capability would be denied where people are deprived of dignity and self-respect.[1136] The principle of each person as an end can thus rephrased into "a principle of each person's capability: the capabilities sought are sought for each and every person."[1137]

DOMINATION AND OPPRESSION - In order to lead a dignified life, people must have access to the necessary conditions that allow them to freely exercise their rights towards self-determination (or self-definition) and self-development. A dignified person holds control over their own life and is not passively shaped or pushed around.[1138] Given the central importance of freedom and choice, the capability approach is also sensitive towards the conditions under which a person's preferences and desires are formed and developed.[1139] For example, preferences can be shaped

---

[1134] Nussbaum, *Creating Capabilities: The Human Development Approach*, 30–31; Claassen, 'Human Dignity in the Capability Approach'.

[1135] Respect is the "proper attitude toward dignity", and together both notions form a "concept-family" that gives expression to the Kantian ideal that a person should be treated as an end, rather than a means. Nussbaum, 'Human Dignity and Political Entitlements', 353–54. Claassen, 'Human Dignity in the Capability Approach', 242.

[1136] According to Nussbaum, equality is a foundational, integral and essential component to the idea of dignity: "it is the *equa*l dignity of human beings that demands recognition," and equality must be added "to the bare idea of dignity" if we want to specify the social goal that is equality of capability in any meaningful way. Nussbaum, *Frontiers of Justice: Disability, Nationality, Species Membership*, 2006, 292. Nussbaum, 'Human Dignity and Political Entitlements', 363; Nussbaum, *The Cosmopolitan Tradition: A Noble But Flawed Ideal*, 91. Nussbaum, *Frontiers of Justice: Disability, Nationality, Species Membership*, 2006, 292–284.

[1137] This also means that equality of capability is not sought in the first place for groups or bodies. These can be important for human flourishing, and therefore in need of support, "but it is because of what they do for people that they are so worthy, and the political goal Is always the promotion of the capabilities of *each person*."Nussbaum, *Women and Human Development: The Capabilities Approach*, 74.

[1138] Nussbaum, 'Women and Equality', 234; Nussbaum, *Women and Human Development: The Capabilities Approach*, 72–73.

[1139] In this context, data protection authorities should closely monitor the obligations decision-makers have under privacy and data protection law to provide meaningful information, that is information that is among others

by tradition and intimidation, or people might adapt their desires based on their environmental conditions.[1140] Likewise, persons might feel pressured to behave or act in a certain way because they know this would result in a preferred outcome.[1141] In case a society wants to hold people responsible for their own choices, including the inequalities they incur as a result of those choices, it should guarantee that these choices can be exercised in a free and informed manner.[1142] Furthermore, the most important powers of humans, including their capacity for choice, are vulnerable and therefore necessitates the existence of support mechanisms that enable and promote their flourishing.[1143]

It would appear then that a capability approach towards justice also opposes itself against a variety of socio-relational (and socio-economic) harms imposed through algorithms. Algorithmic structures and acts of oppression and domination are external constraints that limit a person (and group)'s opportunities to choose and act in socially meaningful and recognised ways. Algorithmic acts of domination limit agency, as they enable others to arbitrarily interfere with a person's choice over options. People can be steered to act in accordance with the goals of the decision-maker. In this sense, people become tools, rather than ends, and as a consequence, denied the respect they deserve. This also holds true for algorithmic acts of oppression, such as the exploitation of decision-subjects, whose actions, behaviours, and preferences can be monitored, analysed, and extracted to the purposes and benefit of others. Algorithmic classification acts can constrain a person's identity formation, lowering the sense of self-worth and self-confidence they need to exert self-definition and self-presentation.[1144] Conversely, decision-subjects should be given enhanced control and agency *inter alia* mechanisms of accountability, transparency,

---

transparent, clear and intelligible, and which allows people to reasonably foresee and assess with a sufficient degree of knowledge the environment and decisions they enter into, including, but not limited to, the inequalities they may incur as a result thereof. Likewise, the capability theory would require citizens to receive education and digital literacy skills concerning the dangers and harms of the online environment and to identify where a choice is either freely given or imposed, and what rights citizens have. See for example on merits of transparency in data protection laws: Naudts, Dewitte, and Ausloos, 'Meaningful Transparency through Data Rights: A Multidimensional Analysis'.

[1140] Nussbaum, *Women and Human Development: The Capabilities Approach*, chaps 2, Adaptive Preferences and Women's Options. On adaptive preferences, see among others: Jon Elster, *Sour Grapes: Studies in the Subversion of Rationality*, Cambridge Philosophy Classics (Cambridge: Cambridge University Press, 2016), https://doi.org/10.1017/CBO9781316494172.

[1141] Wachter, 'The Theory of Artificial Immutability', 17.

[1142] Robeyns, 'The Capability Approach'.

[1143] Nussbaum, 'The Capabilities Approach and the History of Philosophy', 20. Martha C Nussbaum, 'The Supreme Court 2006 Term. Foreword: Constitutions and Capabilities: "Perceptions" Against Lofty Formalism.', *Harvard Law Review* 121 (2006): 11.

[1144] Rawls, *A Theory of Justice*, 1971, 440.

inclusion and (democratic) participation as to reduce their vulnerabilities and remove barriers to their opportunity to select.[1145]

7.3.3 The Central Human Capabilities

People have a fundamental interest to lead a dignified life. In order to lead such a life, Nussbaum believes it is necessary for people to have access to a set of central or core human (combined) capabilities, i.e. real opportunities or substantive freedoms that make available "truly human functioning".[1146] To not have the appropriate social, political, familial, economic and socio-technical conditions in place for people to function in accordance to their (internal) capabilities, is to deny and violate their rightful claim towards human dignity and to deprive them of self-respect.[1147] For this reason, the structure of social and political institutions must be chosen "with a view to promoting at least a threshold level of [the central]] human capabilities." [1148] Because

---

[1145] Because equality of capability should be secured for *each person*, structures of oppression and domination are not the evaluative focal point. As Deneulin observes: "What matters is not what a structure or group is doing, but how each individual is doing and the impact of these structures and groups on the lives of each individual. Deneulin, 'Recovering Nussbaum's Aristotelian Roots', 626. Group-based measures, such as positive obligations (or affirmative actions) can be an effective instrument in raising a person's individual capabilities. Though it is only because such a measure is to the benefit of each and every person that it can be justified. The interests of a group cannot be placed above the interests of an individual, and an individual cannot be asked to sacrifice themselves in favour of a policy that benefits the group they are a member of. At the same time, to assure that every person has an equal entitlement to freedom, may necessitate a society to curtail the freedoms of other individuals and groups. For example, Nussbaum notes that "gender justice cannot be successfully pursued without limiting male freedom Nussbaum, *Creating Capabilities: The Human Development Approach*, 72. Nussbaum, 31. Nussbaum, 72.

[1146] They are a support system that enable each person to flourish, and the capabilities to which they are entitled are "many, and not one;" not only material, but also immaterial. Nussbaum, *Frontiers of Justice: Disability, Nationality, Species Membership*, 2006, 74.; Nussbaum, *Women and Human Development: The Capabilities Approach*, 70–77; Nussbaum, *Frontiers of Justice: Disability, Nationality, Species Membership*, 2006, 159–60; Claassen, 'Human Dignity in the Capability Approach'; John M. Alexander, 'Non-Reductionist Naturalism: Nussbaum between Aristotle and Hume', *Res Publica* 11, no. 2 (June 2005): 157–83, http://dx.doi.org.kuleuven.ezproxy.kuleuven.be/10.1007/s11158-005-1483-6; Alexander, *Capabilities and Social Justice: The Political Philosophy of Amartya Sen and Martha Nussbaum*, chaps 6, Aristotle and Nussbaum's Hybrid Theory of Capabilities; Nussbaum, *Creating Capabilities: The Human Development Approach*, 32. Nussbaum, 'Nature, Function, and Capability'.

[1147] Nussbaum, 'Human Dignity and Political Entitlements', 359; Nussbaum, *Creating Capabilities: The Human Development Approach*, 30–31.

[1148] Whereas the list is presented as a list of "combined capabilities", in some cases, a focus on functionings, instead of capabilities can be desirable. For example, Nussbaum herself has argued that in relation to areas that concern human dignity and self-respect, the actual functioning should be the appropriate aim of given policy. This means that self-respect in itself, rather than the opportunity to receive or obtain self-respect, should be guaranteed at all costs. When

they co-order the social world, algorithmic structures and decision-acts can be evaluated in view of their impact on people's central capabilities (or fundamental entitlements).[1149]

The central capabilities are presented as an open-ended list. They set a benchmark or set of political goals that policy makers can use for aspiration, comparison, and to consider what (in)equality entails.[1150] They are the substantial freedoms that a good society should secure for each and every person. Once secured, people are free to exercise them in action.[1151] The central capabilities are as follows[1152]:

1. **Life:** The ability able to live to the end of a human life (in a way that is worth living).
2. **Bodily health:** The ability to have good health, adequate nourishment, and shelter.
3. **Bodily integrity:** The ability to move freely, secured against violence; having opportunities towards sexual satisfaction and choice over reproductive matters.

---

choosing whether one should focus on capabilities or functioning, Nussbaum offers the following rule of thumb: "the more crucial a function is to attaining and maintaining other capabilities, the more entitled we may be to promote the actual functioning in some cases, within limits set by an appropriate respect for citizen's choices." Nussbaum, *Women and Human Development: The Capabilities Approach*, 92. Nussbaum, *Frontiers of Justice: Disability, Nationality, Species Membership*, 2006, 172. See also: Nussbaum, *Creating Capabilities: The Human Development Approach*, 20. On the choice between capabilities and functionings, see also: Ingrid Robeyns, 'The Capability Approach in Practice*', *Journal of Political Philosophy* 14, no. 3 (2006): 353, https://doi.org/10.1111/j.1467-9760.2006.00263.x.

[1149] Martha Nussbaum, 'Capabilities as Fundamental Entitlements: Sen and Social Justice', *Feminist Economics* 9, no. 2–3 (January 2003): 33–59, https://doi.org/10.1080/1354570022000077926.

[1150] Despite people's individualised conceptions of the good, the central capabilities can become the object of an overlapping consensus. Nussbaum, 'Women and Equality', 89; Nussbaum, *Frontiers of Justice: Disability, Nationality, Species Membership*, 2006, 70–74. Alexander, John M., 'Capability Egalitarianism and Moral Selfhood', 7.

[1151] Nussbaum, *Creating Capabilities: The Human Development Approach*, 18.

[1152] These capabilities are presented in (simplified) fashion based on how they last appeared in Nussbaum's (most) recent work on the cosmopolitan tradition. Here, I present a summarised view. For a full overview, see for example her inclusion of the ten capabilities in the following works: Nussbaum, *Women and Human Development: The Capabilities Approach*; Nussbaum, *Frontiers of Justice: Disability, Nationality, Species Membership*, 2006; Nussbaum, *Creating Capabilities: The Human Development Approach*; Nussbaum, *The Cosmopolitan Tradition: A Noble But Flawed Ideal*. For a full version of the central capabilities, reference can be made to the latter works. Since their introduction, Nussbaum has repeatedly stated that these central capabilities must not be seen as fixed. For example, in The Cosmopolitan Tradition she notes "This list, humble and revisable, is an abstract template that can be further specified in accordance with a particular nation's history and material circumstances. It was formulated in part by studying the constitutions of nations such as India and South Africa that have focused intensely on human dignity, and it can be used similarly to ground projects of constitution-making or basic legislation." Nussbaum, 243. In other work, Nussbaum has noted how the list came about through cross-cultural discussion and can always be contested or remade. Nussbaum, 'Women and Equality', 236.

4. **Senses, imagination and thought:** The ability to use one's senses, imagination, mind, and reason, informed and cultivated by adequate education, protected by freedom of expression regarding artistic and political speech, and freedom of religious exercise. People should enjoy the freedom to have pleasurable experiences and be free from non-beneficial pain.
5. **Emotions:** The ability to form attachments to both things and people, free from fear and anxiety, and to support forms of association conducive to their development.
6. **Practical reason:** The ability to form of conception of the good, and to engage in critical reflection about the planning of one's life.
7. **Affiliation:**
    a. The ability to live, recognise and interact with and towards others, through (among others) the protection of affiliation, assembly, and political speech
    b. The ability to access the social bases of self-respect and non-humiliation and to be treated as a dignified being of equal moral worth, through (among others) equality and non-discrimination laws.[1153]
8. **Other species:** The ability to live with concern and relation to animals, plants, and natural world.
9. **Play:** The ability to laugh, play and enjoy recreational activities
10. **Control over one's environment:**
    a. **Political:** The ability to participate into the conditions (political choices) that govern one's life, through freedom of participation, speech, and association.
    b. **Material:** The ability to hold property, to seek and hold work in meaningful relationships of mutual recognition, and to be free from unwarranted search and seizure.

PRIMARY GOODS AND HUMAN RIGHTS? - The central capabilities are a "long list of opportunities for functioning such that it is always rational to want them whatever else one wants."[1154] In this sense, they are similar to the primary goods. Still, there are significant differences between them. The primary goods were social in nature. The central capabilities also include "natural goods", such as health, vigour, intelligence, and imagination.[1155] Furthermore, material goods or "thing-like

[1153] On the basis of race, sex, sexual orientation, ethnicity, caste, religion and national origin. Nussbaum, *The Cosmopolitan Tradition: A Noble But Flawed Ideal*, 242.

[1154] Nussbaum, *Women and Human Development: The Capabilities Approach*, 88.

[1155] For example, by including the senses, imagination and thought, as well as emotions as central capabilities, the list can offer a direct argument against the irresponsible deployment of emotion recognition software or software that

items", such as income and wealth, do not have intrinsic value. They should be looked at for their instrumental worth to develop and exercise the capabilities.[1156] Being focused on real opportunities, the capabilities approach affirms the ideal that in order to put certain oppressed or disadvantaged groups in a similar position of those that are privileged, more resources will need to be spent on the former.[1157]

PRIMARY GOODS AND HUMAN RIGHTS - The capabilities can also be likened to human rights though Nussbaum highlights significant differences.[1158] First, the central capabilities clearly impose an affirmative task towards the securitisation of capabilities and rights, and the removal of (external) obstacles in their enjoyment.[1159] The capabilities approach is positively aimed towards the transformation of society in order to ensure equality of capability for each person. While the central capabilities highlight the importance of non-discrimination laws, to ensure formal equal treatment or to prohibit discrimination on a number of selected grounds, may not be sufficient to equalise the living conditions from all.[1160] Second, the capability approach challenges the oft-held distinction between, or conceptual independence of, the civil and political realm and that of socio-economic arrangements.[1161] Civil and political liberties can only be met in practice under the right socio-economic preconditions. The central capabilities thus capture civil and political, as well as social and economic opportunities. Finally, because people's beings and doings are shaped by people's private or socio-relational circumstances, the capability approach also actively challenges the traditional divide between the public and private realm.[1162]

---

makes uses of surreptitious or manipulative techniques, such as dark patterns. If the deployment of these techniques interferes with people's capabilities, they could be forbidden on those grounds. Nussbaum, 88–89.

[1156] Nussbaum, 89.

[1157] Nussbaum, *Creating Capabilities: The Human Development Approach*, 57.

[1158] Nussbaum, *Frontiers of Justice: Disability, Nationality, Species Membership*, 2006, 284. Nussbaum, 'The Capabilities Approach and the History of Philosophy', 26.

[1159] Martha C. Nussbaum, *Frontiers of Justice: Disability, Nationality, Species Membership* (Cambridge, Massachusetts ; London, England: Belknap Press: Harvard University Press, 2006), 287–88.

[1160] Nussbaum, *Creating Capabilities: The Human Development Approach*, 31.

[1161] Similarly, it questions the priority of first over second generation rights. For example, Nussbaum also notes how Rawls' lexical ordering seems to imply that liberty can exist independently from socio-economic redistribution. Liberties do not exist only on paper but are dependent upon social and material preconditions. If they are to be met in practice. Because every right is dependent upon economic and social preconditions, every right is in some sense a socio-economic right. Nussbaum, *Frontiers of Justice: Disability, Nationality, Species Membership*, 2006, 288–90. Nussbaum, *Creating Capabilities: The Human Development Approach*, 67.

[1162] Nussbaum highlights how women suffer inequality also due to their position in the family, which may deprive them of educational opportunities. Nussbaum, *Frontiers of Justice: Disability, Nationality, Species Membership*, 2006, 290.

SUFFICIENCY OR EQUALITY? – Unlike Rawls' principles of justice, the rules that cover the "distribution" of the central capabilities are less well-defined. The principle rule is that of the threshold: "for each important entitlement, there is some appropriate level beneath which it seems right to say the relevant entitlement has not been secured."[1163] Certain central capabilities, such as the social bases for self-respect, non-humiliation and the political liberties, can only be adequately secured when they are equally secured. For other capabilities, such as those connected to the idea of property or instrumental goods, the appropriate standard is one of adequacy or sufficiency. [1164] In order to establish what standard would be appropriate, Nussbaum believes dignity can be used as a benchmark: "Some capabilities must be secured to citizens on a basis of equality, or equal dignity has not been respected. Others do not seem to have this intrinsic relationship to dignity; with these the capabilities approach supplies a threshold of adequacy."[1165] Likewise, in some cases, it is not the capability, but the function that should be prioritised. Governments should promote policies that treat people with actual dignity and respect. It is insufficient to give people the option to be treated with dignity. Functions are a legitimate a policy goal when they appear essential for the realisation and preservation of other capabilities [1166]

7.3.4 The Central Capabilities: An Evaluative Benchmark for Algorithmically Guided Decision-Making

Though one could focus on the threshold, Coeckelbergh believes the central capabilities could not only be seen as minimal requirements for justice and dignity, but also as a formulation of an ethical

[1163] Nussbaum, 291–92.

[1164] For instance, justice requires not for every person to have an equal house. Adequate housing for everyone seems to appropriate standard to attain. Nussbaum, *Frontiers of Justice: Disability, Nationality, Species Membership*, 2006, 293. Elsewhere, Nussbaum has said: "To aim at complete equality [may] fetishize material goods too much, and a generous threshold seems a reasonable target." Nussbaum, *The Cosmopolitan Tradition: A Noble But Flawed Ideal*, 244.

[1165] Conversely, in these instances (material or strict) equality need be needed in order to ensure human dignity or equal human dignity. While people may feel envious in the face of socio-economic inequality, these feelings can sometimes be ascribed to people's excessive valuation or fetishisation of material goods, rather than a violation of their dignity. Societies may differ in their valuation of these inequalities, and they may thus settle on different distributive rules. However, if these (material) things are valuated more highly due to a social norm, a society does not stop to be just if it chooses not to honour this valuation. Nussbaum, *Frontiers of Justice: Disability, Nationality, Species Membership*, 2006, 293–95.

[1166] Nussbaum, *Women and Human Development: The Capabilities Approach*, 91–92.

maximum: it identifies those elements that make the good life and not only human life possible.[1167] In this context, Nussbaum has argued that justice should maintain its focus on the outcome**:** every procedure that succeeds in achieving the outcome is a good or just procedure.[1168] Following the aforementioned, the central capabilities may not only be a source of political or regulatory principles, but also an interpretative instrument that can inform decision-makers in their actions.[1169] More specifically, the central capabilities can act as a concrete list of criteria through which algorithmic decision-making structures and systems can be evaluated in terms of their contribution to, enhancement or limitation of people's enjoyment of them.[1170] Within the context of the capabilities approach, algorithmic decision-making structures and decisions are best seen as external or environmental factors that are part of people's social, economic, political and socio-technical living environment and can affect people in their substantial freedoms to achieve potential beings and doings. Following Nussbaum, the central capabilities can tell us "what to consider salient" within the algorithmic governance debate and not "dictate a final assignment of weights and a sharp-edged decision."[1171]

In a negative sense, the list can be used to assess the potential harmful impact of algorithmically guided decision-making structures and processes: do they negatively interfere with people's central capabilities, and if so, could the decision-maker justify this interference? In a positive sense, the central capabilities function as a guide and indicate those substantive freedoms decision-makers should maintain, restore, or perhaps enhance.[1172] For instance, social networks could be evaluated in light of their contribution to people's capacity for affiliation, and their capacity to engage in socially meaningful relationships: do they generate relationships of mutual respect or hostility? [1173] Opaque recommender systems reduce decision-subjects agency and

---

[1167] Mark Coeckelbergh, 'Human Development or Human Enhancement? A Methodological Reflection on Capabilities and the Evaluation of Information Technologies', *Ethics and Information Technology* 13, no. 2 (1 June 2011): 81–92, https://doi.org/10.1007/s10676-010-9231-9.

[1168] Nussbaum, *Frontiers of Justice: Disability, Nationality, Species Membership*, 2006, 82. In this regard, it could be considered "a cousin of consequentialism." Nussbaum, *Creating Capabilities: The Human Development Approach*, 95.

[1169] Nussbaum, *Creating Capabilities: The Human Development Approach*, 74–75.

[1170] Mark Coeckelbergh, 'Health Care, Capabilities, and AI Assistive Technologies', *Ethical Theory and Moral Practice* 13, no. 2 (1 April 2010): 185, https://doi.org/10.1007/s10677-009-9186-2.

[1171] Nussbaum, *Creating Capabilities: The Human Development Approach*, 41.

[1172] Coeckelbergh, 'Health Care, Capabilities, and AI Assistive Technologies', 185.See also: Jason Borenstein and Yvette Pearson, 'Robot Caregivers: Harbingers of Expanded Freedom for All?', *Ethics and Information Technology* 12, no. 3 (1 September 2010): 277–88, https://doi.org/10.1007/s10676-010-9236-4.

[1173] Coeckelbergh, 'Human Development or Human Enhancement?', 1 June 2011, 88.

choice.[1174] Yet, by giving decision-subjects choice and control over their influence and operations, recommenders could be positively harnessed to facilitate decision-subjects in their navigation of the digital environment. Doing so could maximise user freedom and choice, by selecting and identifying those options that are most strongly associated to a person's actual interests.[1175]

Whereas the abovementioned examples highlight the potential harms and contributions of algorithms more generally, in a similar vein, the list of central capabilities may function as a benchmark to evaluate algorithmic classification and differentiation acts, and the (in)equalities they produce, whether they be socio-relational, socio-economic, or otherwise related to people's equal access to the central capabilities. Under this view, algorithmically generated inequalities will be more difficult to justify where they diminish a person or group's substantial freedoms and opportunities. For instance, Obermeyer and others discovered how certain commercial care algorithms systematically privilege white patients over black patients.[1176] White patients were allowed to cut in line for special programs related to complex and chronic health conditions. Even if they had the same level of need, due to racial bias, black people were more likely to be excluded from extra care programs.[1177] This bias occurred because the algorithms used health costs as a proxy for health needs. The use of health costs appeared both convenient and effective. Due to social conditions, black patients have uneven access to, and as a consequence, can spend less money on health care. Even when they had the same level of need, the algorithm falsely concluded that Black patients were typically healthier than equally sick white patients. Not only is this inequality incompatible with the requirement that people ought to be free from discrimination on the basis of ethnicity, but the system also denied black patients in their access to the central capabilities of life, bodily health, and bodily integrity. Under the capability approach, an active duty exists to address the capability failures that are the result of underrepresentation. Decision-makers should have invested into research that would have improved the reliability and predictive quality of their data, rather than settle for convenience. At the same time, public institutions should have made better efforts to equalise access to health care.[1178]

[1174] Alan Lundgard, 'Measuring Justice in Machine Learning', *Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency*, 27 January 2020, 25, https://doi.org/10.1145/3351095.3372838.

[1175] Danaher, 'Freedom in an Age of Algocracy', 267.

[1176] Ziad Obermeyer et al., 'Dissecting Racial Bias in an Algorithm Used to Manage the Health of Populations', *Science* 366, no. 6464 (25 October 2019): 447–53, https://doi.org/10.1126/science.aax2342; Tom Simonite, 'A Health Care Algorithm Offered Less Care to Black Patients', *Wired*, accessed 26 April 2022, https://www.wired.com/story/how-algorithm-favored-whites-over-blacks-health-care/.

[1177] Obermeyer et al., 'Dissecting Racial Bias in an Algorithm Used to Manage the Health of Populations'; Simonite, 'A Health Care Algorithm Offered Less Care to Black Patients'.

[1178] Obermeyer et al., 'Dissecting Racial Bias in an Algorithm Used to Manage the Health of Populations'.

Algorithmically generated inequalities could find justification if they were to enhance or improve people in their access to the central capabilities.[1179] Care algorithms could justifiably differentiate between patients so to better meet each person's health needs. In France, the benefits of algorithms were explored as a tool to combat favouritism in day care allocation.[1180] Given the shortage of day care places, the practice of cronyism is particularly problematic. In one city, the distribution of day care slots was automated. Among families who had the same priority score, a random choice was made. This use of automated decisions was also made transparent to families. Though the underlying shortage problem was not resolved, the system did try to increase the fairness of day car allocation by aligning the distribution of spaces with the needs and wishes of parents.[1181] The system was inspired by school matching algorithms. These algorithms try to lower the risk for a student to be rejected from their preferred school of choice, while other students who had assigned a lower preference score to that school would be matched to that school.[1182] Because these algorithms try to differentiate based on criteria that try to connect people to their actual choice, they could be seen as agency and freedom enhancing.

REALMS OF OPPORTUNITY - Against this backdrop, it could be argued that the central capabilities, due to their expansive nature, impose a duty onto decision-makers (and regulators) to contemplate the multiple ways through which algorithms can interfere with people's lives and opportunities, and how their impact might take place over extended periods of time and across multiple areas of life.[1183] As aforementioned, certain algorithmic decisions, such as hiring or lending decisions, govern so-called "realms of opportunity."[1184] In this context, the central capabilities could function as a yardstick and indicator of power. The more a decision impacts

---

[1179] Luiz Costa, *Virtuality and Capabilities in a World of Ambient Intelligence*, vol. 32, Law, Governance and Technology Series (Cham: Springer International Publishing, 2016), https://doi.org/10.1007/978-3-319-39198-4.

[1180] 'In French Daycare, Algorithms Attempt to Fight Cronyism', *AlgorithmWatch* (blog), accessed 26 April 2022, https://algorithmwatch.org/en/algorithms-to-fight-cronyism-in-french-daycare/. Thibaut Schepman, 'Gros pistons et petites manip pour une place en crèche à Boulogne-Billancourt', BuzzFeed, accessed 3 October 2022, https://www.buzzfeed.com/fr/thibautschepman/gros-pistons-et-petites-manip-pour-une-place-en-creche-a-bou. Alexandre Léchenet, 'A Valence-Romans, des places en crèche attribuées par un algorithme', La Gazette des Communes, 20 August 2020, https://www.lagazettedescommunes.com/691390/a-valence-romans-des-places-en-creche-attribuees-par-un-algorithme/.

[1181] 'In French Daycare, Algorithms Attempt to Fight Cronyism'. Yuichiro Kamada and Fuhito Kojima, 'Fair Matching under Constraints: Theory and Applications', n.d., 55.

[1182] 'In French Daycare, Algorithms Attempt to Fight Cronyism'.

[1183] See also Chapter 1, Section 1.2.2. "Scale and Amplification", the negative spiral argument. Citron and Pasquale, 'The Scored Society: Due Process for Automated Predictions', 2014; Zarsky, 'Understanding Discrimination in the Scored Society'.

[1184] Creel and Hellman, 'The Algorithmic Leviathan', 2021, 37–38.

people in their access to a range of central capabilities, the more significance its interference in their freedom to choose between alternative combinations of functionings. For instance, a person's ability to land a steady and meaningful job gives a person a sense of self-worth and self-respect and expands their choice when planning their life. In instances like these, decision-makers ought to advance weighty reasons in order to motivate why, in light of the interests they pursued, differentiation, and the associated socio-relational and socio-economic inequalities decision-subjects may incur, was justified. Yet this egalitarian disadvantage should not only be framed in relation to the particular decision at hand, but in terms of the greater exclusionary harm decision-subjects risk facing (over time) in other areas of life. Regulators could oblige decision-makers to identify as part of their duties fertile advantages and corrosive disadvantages their decisions might yield.[1185] Fertile advantages are likely to enhance or promote other (related) capabilities. Corrosive disadvantages do the opposite. These monitoring obligations would motivate decision-makers to take responsibility for their position of power. They should actively look outward and contemplate the wider impact systems can have once they are placed within a larger socio-technical, political, and economic environment by considering *among others* its interconnected nature and the unique amplification properties of algorithms.[1186]

Of course, algorithmic decision-makers have their own freedoms and interests to pursue. A company cannot hire everyone, nor can a financial institution grant every person a loan. Still, decision-makers should weigh their interests against the capability interests of decision-subjects. Though they might not have an active duty to maximise, or enhance, decision-subjects' capabilities, they should still respect their fundamental interests. Still, even in these cases, regulators can give decision-subjects the resources and protection they need to maximise their freedoms, choice, and knowledge within the algorithmic environment. For instance, the education of digital literacy, combined with a strong set of informational rights, could help decision-subjects to better assess the long-lasting impact their data and data-driven decisions might have on their lives.

[1185] These two notions (fertile and corrosive disadvantage) were introduced by Wolff and De-Shalit, but adopted by Nussbaum. Nussbaum, *Creating Capabilities: The Human Development Approach*, 144–45. Nussbaum, 100. Nussbaum observes how corrosive disadvantages will often "be failures connected to marginalisation, stigma, and other forms of group-based powerlessness, giving societies reason to adopt group-based remedies, even though the end in view is always the full empowerment of each individual." Jonathan Wolff and Avner de-Shalit, *Disadvantage*, *Disadvantage* (Oxford University Press, 2007), https://www.oxfordscholarship.com/view/10.1093/acprof:oso/9780199278268.001.0001/acprof-9780199278268; Jonathan Wolff and Avner de-Shalit, 'On Fertile Functionings: A Response to Martha Nussbaum', *Journal of Human Development and Capabilities* 14, no. 1 (February 2013): 161–65, https://doi.org/10.1080/19452829.2013.762177.

[1186] See Chapter 1 Section Scale and Amplification.

ARCHITECTONIC CAPABILITIES – Among the capabilities listed, those of affiliation and practical reason hold a special value. According to Nussbaum, these two capabilities, "play a distinctive architectonic role"[1187] The capability of practical reason proclaims the centrality of personal choice, or agency, holds in the conception of capability as freedom.[1188] Practical reason represents a person's capacity to plan one's life in accordance with their conception of the good.[1189] Whereas practical reason affirms people as life makers or self-authors, affiliation affirms them as social beings.[1190] People should be able to engage in various forms of social interaction, and be able to imagine the lived experiences of others.[1191] At the same time, affiliation grants people the social bases for self-respect and non-humiliation, and a right to be recognised as a beings with equal moral worth.[1192] The central capability of affiliation thus necessitates that sufficient attention is given to the relationships that people may hold, whether they are familial, group-based, political, or algorithmically generated.[1193] The capabilities of practical reason and affiliation can be marked out as preconditions for people to develop and exercise their right to self-determination and self-development. Their foundational importance re-affirms the aforementioned belief that algorithmic structures and acts of domination and oppression, including unjustified generalisations, should be seen as particularly corrosive. As enemies towards self-determination and self-development, oppression, and domination close of people's access to the capability of practical reason and affiliation, and in turn, their access to the other capabilities, making them more vulnerable in the process.

[1187] Nussbaum, *Creating Capabilities: The Human Development Approach*, 39.

[1188] Nussbaum, 39.

[1189] Nussbaum, 39. Together with human dignity, practical reason may be considered the core guiding idea underlying the CA: people shape and make their lives and they should be treated as an end, and never as solely as an instrument for the ends of others. Nussbaum, *Women and Human Development: The Capabilities Approach*, 284.

[1190] Nussbaum, *Creating Capabilities: The Human Development Approach*, 39.

[1191] Whereas the capability approach shows relational concerns, it nonetheless remains distributional in nature in the sense that the central capabilities are about the things people are able to be and do. That is, it remains focused on achieving a specific outcome, e.g. give people the opportunity to enter into meaningful relationships, rather than impose onto people a duty to engage in ways that are mutually respectful. Baker, 'Conceptions and Dimensions of Social Equality', 76–77.

[1192] Nussbaum, *Creating Capabilities: The Human Development Approach*, 34.

[1193] Nussbaum, 39–40.

7.3.5 The Central Capabilities: A Transformative Approach towards Algorithmic Decision-Making?

A capability approach towards justice requires decision-makers and regulators to consider the dynamic or diverse nature of human capabilities. The capability approach towards justice wants to place *every* person in a position that is worthy of living.[1194] One of the reasons why people fail to enjoy substantial freedom is because they face discrimination and marginalisation, as well as systemic and entrenched inequalities.[1195] Similarly, people can differ greatly in their need for resources, and they can depend upon others to satisfy those needs.[1196]

In order to achieve true equality of opportunity, freedom and choice, regulators should address the various socio-relational and socio-economic inequalities that limit persons or group's in their access to the central capabilities.[1197] In certain domains of life, those with less advantage may therefore need to be given more of something in order to reach the desired threshold level of a given capability.[1198] This however, might require policy makers to curtail the freedoms of others .[1199] For example, algorithmic decision-makers should be curtailed in their freedom to oppress and dominate decision-subjects.

In order to overturn the inequalities and the lack of liberties and opportunities certain persons or groups experience within specific decision-making environments, decision-subjects might need to be treated differently as to accommodate their needs. Because the capability approach demands the (rough) equalisation of substantial freedoms, it also mandates transformation. Within a society characterised by structures of inequality, it must be ascertained to what extent algorithms can be an active driver for change, built with the specific purpose of dismantling, rather than introducing the entrenched social injustice faced by certain members and groups within society.[1200] In this context, algorithmic decision-makers might not suffice with the mere

---

[1194] Nussbaum, *Women and Human Development: The Capabilities Approach*, 74; Nussbaum, 'Women and Equality', 233.

[1195] Nussbaum, *Creating Capabilities: The Human Development Approach*, 19.

[1196] Nussbaum, *Frontiers of Justice: Disability, Nationality, Species Membership*, 2006, 87–89.

[1197] Schuppert, 'Suffering from Social Inequality', 108.

[1198] On the role of emotions, such as compassion and solidarity, see for example: Nussbaum, *Creating Capabilities: The Human Development Approach*, 180–84.

[1199] For example, Nussbaum notes that "gender justice cannot be successfully pursued without limiting male freedom." Nussbaum, 72.

[1200] Eubanks, 'The Digital Poorhouse', para. 17. As also referenced by: Hoffmann, 'Where Fairness Fails', 907. In this context, the capability approach aligns with substantive approaches towards equality. See for example, the transformative dimension in Fredman's account of substantive equality. Fredman, 'Substantive Equality Revisited',

implementation of "avoidance" strategies. For instance, a common policy solution is to "debias" datasets. Following Balayn and Gürses, these strategies concern "the application of select methods to address bias by achieving certain forms of statistical parity (e.g. making sure that the accuracy of a recidivism prediction system is similar for Black and White people by rebalancing a training dataset and re-training a machine learning model)."[1201] De-biasing strategies are developed in an effort to reduce the risk towards discrimination and the perpetuation of historic injustice. In practice, these techniques might only mitigate, social injustice and not correct it.[1202] By focusing only on the data that is used within a specific decision-making context, one is more likely to maintain the status quo, rather than secure each person equality of capability. In order to achieve the latter, these strategies should be accompanied by measures aimed to remove additional institutional and structural barriers towards equality.[1203] The capability approach is therefore demanding in its goal to ensure for each person equality of capability.[1204]

How then to align algorithmic decision-making structures and acts with these transformative goals? In this context, recourse could be made to conceptualisations of social equality as a positive ideal. Models of participatory and democratic equality, which focus on values like inclusion and public accountability, could ensure that decision-making systems are built with the needs and

729–30. Khaitan too has argued that the "primary purpose of discrimination law is to secure to us an aspect of freedom that is essential in order to live a flourishing life." Tarunabh Khaitan, 'The Point of Discrimination Law: Securing the Freedom to Flourish', in *The Empire of Disgust: Orejudice , Discrimination, and Policy in India and the US*, ed. Zoya Hasan et al. (Delhi: Oxford University Press, 2018), 366, https://doi.org/10.1093/oso/9780199487837.003.0017. See also: Dorian R. Woods, Yvonne Benschop, and Marieke van den Brink, 'What Is Intersectional Equality? A Definition and Goal of Equality for Organizations', *Gender, Work & Organization* 29, no. 1 (2022): 92–109, https://doi.org/10.1111/gwao.12760. See also: Martha C. Nussbaum, 'Ageing, Stigma, and Disgust', in *The Empire of Disgust* (Delhi: Oxford University Press, 2018), 166, https://doi.org/10.1093/oso/9780199487837.003.0008.

[1201] Balayn and Gurses, 'Beyond Debiasing: Regulating AI and Its Inequalities.', 23.

[1202] For more information on debiasing as a techno-centric (policy) approach towards addressing inequality, and its limitations and flaws, see also: Balayn and Gurses, 23.

[1203] See for instance the work of Wachter and others on "bias transforming" fairness metrics within machine learning: Sandra Wachter, Brent Mittelstadt, and Chris Russell, 'Bias Preservation in Machine Learning: The Legality of Fairness Metrics Under EU Non-Discrimination Law', SSRN Scholarly Paper (Rochester, NY: Social Science Research Network, 15 January 2021), https://doi.org/10.2139/ssrn.3792772. Balayn and Gurses, 'Beyond Debiasing: Regulating AI and Its Inequalities.'

[1204] Whereas the capability approach is transformative in the sense that it looks to provide real opportunities for everyone, a next step could be to enhance or expand upon people's overall freedoms through technological means. Ilse Oosterlaken, 'Design for Development: A Capability Approach', *Design Issues* 25, no. 4 (October 2009): 91–102, https://doi.org/10.1162/desi.2009.25.4.91. Mark Coeckelbergh, 'Human Development or Human Enhancement? A Methodological Reflection on Capabilities and the Evaluation of Information Technologies', *Ethics and Information Technology* 13, no. 2 (June 2011): 81–92, https://doi.org/10.1007/s10676-010-9231-9.

freedom of those affected in mind.[1205] For example, inclusive, co- or participatory design strategies see decision-subjects as partners and experts of their own experience.[1206] Such approaches ask decision-makers to take into account the lived experiences of decision-subjects and to learn and understand how a decision can affect their substantial freedoms.[1207] For example, drawing on the work of Murphy and Gardoni, Lundgard proposes a multi-stage, participatory and democratic approach in order to take people's capabilities into account during the design of a decision-making system. [1208] Under a first step, decision-makers would collaborate with decision-subjects whose well-being might be affected by a specific system, and select what substantive freedoms are at stake. Once the relevant freedoms have been identified (and formalised), the system would be built in such a way that all affected parties are brought to a baseline threshold level of capability achievement.[1209] Furthermore, through their engagement, decision-subjects are given access to the central capability of control over one's environment as they would have the possibility to participate in the (political) choices that govern their lives, which in turn, helps to maximise their freedom and choice within the algorithmic realm (and beyond). These democratic approaches are demanding and time-consuming and perhaps not appropriate for every algorithmically guided decision-making process. Still, where decision-makers hold a position of power, such as when they govern a realm of opportunity, they could be considered an additional safeguard to protect people's substantial freedoms.

[1205] Young, *Inclusion and Democracy*, 23–26.

[1206] See for example: Froukje Sleeswijk Visser et al., 'Contextmapping: Experiences from Practice', *CoDesign* 1, no. 2 (1 April 2005): 119, https://doi.org/10.1080/15710880500135987; Elizabeth B.-N. Sanders and Pieter Jan Stappers, 'Co-Creation and the New Landscapes of Design', *CoDesign* 4, no. 1 (1 March 2008): 8 and 12, https://doi.org/10.1080/15710880701875068; Maarten Van Mechelen et al., 'Collaborative Design Thinking (CoDeT): A Co-Design Approach for High Child-to-Adult Ratios', *International Journal of Human-Computer Studies* 130 (October 2019): 179–95, https://doi.org/10.1016/j.ijhcs.2019.06.013; Maarten Van Mechelen et al., 'The GLID Method: Moving from Design Features to Underlying Values in Co-Design', *International Journal of Human Computer Studies* 97 (2017): 116–28, https://doi.org/10.1016/j.ijhcs.2016.09.005; Oscar Alvarado et al., 'Foregrounding Algorithms: Preparing Users for Co-Design with Sensitizing Activities', 2020, 7. Harambam, Helberger, and van Hoboken, 'Democratizing Algorithmic News Recommenders'.

[1207] Elizabeth Bondi et al., 'Envisioning Communities: A Participatory Approach Towards AI for Social Good', in *Proceedings of the 2021 AAAI/ACM Conference on AI, Ethics, and Society*, 2021, 425–36, https://doi.org/10.1145/3461702.3462612.

[1208] Lundgard, 'Measuring Justice in Machine Learning'. See also for an early mention of the approach in light of machine learning: Gajane and Pechenizkiy, 'On Formalizing Fairness in Prediction with Machine Learning'. Colleen Murphy and Paolo Gardoni, 'Design, Risk and Capabilities', in *The Capability Approach, Technology and Design*, ed. Ilse Oosterlaken and Jeroen van den Hoven, Philosophy of Engineering and Technology (Dordrecht: Springer Netherlands, 2012), 173–88, https://doi.org/10.1007/978-94-007-3879-9_10.

[1209] Lundgard, 'Measuring Justice in Machine Learning', 22–23.

INDIVIDUAL JUSTICE? - The capability approach wants to "produce capabilities for each and every person, and not to use some people as a means to the capabilities of others or of the whole."[1210] This focus also translates to the policies that should be adopted. Group-based measures, such as positive obligations (or affirmative actions) can be an effective instrument in raising a person's individual capabilities. Yet, it is only because such a measure is to the benefit of each and every person that it can be justified.[1211] The interests of a given group cannot be placed above the interests of group members, and group members cannot be asked to sacrifice themselves in favour of a policy that benefits the overall group.[1212] Yet, given the fact that algorithms often rely and apply group-level information, rather than individual insights, hey might lack the specificity to properly respect the diversity of human needs, algorithms might fail to provide individual justice.[1213]

Consider again the central capability of control over one's environment. This central capability states that every person should be able to have the right to seek employment on an equal basis. Suppose the government wants to build an automated decision-making system that will help its unemployed citizens to find their way back on the job market. Such a system would be flawed if it would give to every unemployed citizen an equal amount of aid. Suppose the government wants to build a ranking system which will classify job seekers into distinct categories. The system is based upon clear and reliable statistical unemployment data. It considers a variety of factors that are likely to contribute to a person's unemployment status, including institutional barriers that cause certain persons or groups to face greater disadvantage or adversity on the job market. The government also performs a study which identifies people with an ethnic background or disability as more likely to remain inactive or unemployed.[1214] The government further strengthens the system through a variety of consultation rounds and participatory design strategies, trying to gain

---

[1210] Nussbaum, *Creating Capabilities: The Human Development Approach*, 35.

[1211] Nussbaum, 31. Of course, this does not mean that people cannot identify themselves with certain groups or collectives. Yet, a policy should focus on what it can do for the individuals within that group. Nussbaum's main concern would be the adoption of a policy which favours the group above the interests of the individuals that comprise the group.

[1212] Similarly, social structures are not the evaluative focal point, but the way these structures and affect the experiences of individuals. Deneulin, 'Recovering Nussbaum's Aristotelian Roots', 626.

[1213] Viljoen, 'A Relational Theory of Data Governance'. Reuben Binns, 'On the Apparent Conflict Between Individual and Group Fairness', *ArXiv:1912.06883 [Cs, Stat]*, 14 December 2019, http://arxiv.org/abs/1912.06883.

[1214] See for example: 'New ILO Database Highlights Labour Market Challenges of Persons with Disabilities', ILOSTAT (blog), 13 June 2022, https://ilostat.ilo.org/new-ilo-database-highlights-labour-market-challenges-of-persons-with-disabilities/. Eugenie D'Hooghe, 'Ook in Brussel minder kans bij sollicitaties voor mensen met migratieroots', De Morgen, 20 March 2021, https://www.demorgen.be/gs-bd34de11.

insight into the lived realities of those affected by the system, including but not limited to the aforementioned groups of persons. The system is designed and will use several categories to rank job seekers in light of their specific needs. Based on the category they are assigned the system will propose a number of resources for aid job seekers can have access to. Having filled in an online questionnaire, people who identify as having a disability and ethnic background will be ranked in the category that is eligible to obtain the most aid. They are given personal assistance when looking for jobs, as well as the opportunity to enrol into a number of classes to acquire higher qualifications. Such an approach towards algorithmic decision-making seems beneficial and appears in correspondence with the capability approach. The government clearly wants to build a system that equalises the chances of those identified as having the greatest needs, and it tries to break down other institutional barriers that might have been limiting, such as uneven access to education.[1215] Still, the end-goal is to secure for *each person* equality of capability.

Individualised justice can be guaranteed by complementing the algorithm's group-level reasoning with individual mechanisms of redress, such as having a human in the loop, a general right towards contestability, or a right to an explanation. Such additional safeguards could serve multiple functions. First, algorithmic categorisations tend to be over- and under-exclusive, thereby increasing the risk of false positives and negatives. As a consequence, some people deserving of a given benefit, may be deprived of one (or vice versa). Individual review mechanisms can enable human case workers to identify those instances where a benefit was unjustifiably allocated. Those resources could then be repurposed to benefit those with greater needs. At the same time, decision-subjects are given the opportunity to have individualised information considered when they feel a decision made about them was wrong.[1216] Second, broad categorisations might lack the level of granularity needed to accommodate the differing needs of the various group members. People who have different disabilities might need catered help. Finally, and as above-mentioned, certain classification acts could negatively impact people's sense of confidence and self-respect.[1217] This might be particularly the case where people are labelled as a member of a vulnerable group, or otherwise attached a label that can be perceived as negative, such as being described as having difficulty re-entering the job-market. A purely automated

[1215] It is interesting to note that the Austrian AMS system was actually committed in spending more resources in the support of women. Moreover, the decision's ranking system was built around people's prospects of re-entering the labour market. The CA would also require developers to adequately take into account the cultural and societal specificities of the environment in which their decision-making system would be deployed. In country A, the job market may be most hostile towards women or people with a disability, whereas in country B most challenges may be faced by people with a multi-ethnic background.

[1216] See Chapter 6, Section 6.2.2 The Problematic Nature of Generalisations

[1217] See above, Section 7.2.2 The Social Bases of Self-Respect.

response could be experienced as dehumanising or deindividualizing. In these situations, the presence of a human decision-maker could be more positively received. It gives decision-subjects the sense that they are heard and recognised as individuals, and not solely as a member of a group.[1218] In so doing, algorithmic decision-making processes could address the disadvantage a person might experience due to their group membership. The presence of individual mechanisms could ensure that the unique individuality of each person is respected and the specific needs they have *as* individuals are addressed. Of course, these individualised mechanisms can only function where the humans behind the machine are free prejudicial dispositions as well.

7.4 Conclusion

In this Chapter, the role of distributive conceptualisations of equality for the regulation of algorithmically guided decision-making processes was explored. Algorithms mediate and co-shape the basic institutions of society. They co-govern the distribution of certain justice-relevant goods, the access to which should be (roughly) equal for all.[1219] Against this backdrop, the outcomes of algorithmic decision-making structures and acts can be evaluated in terms of their impact on people's distributive shares. Algorithmic decision-making structures and acts, including the socio-relational and socio-economic inequalities they generate, will produce allocative harms when they fail to guarantee the desired distribution of justice-relevant goods. These acts can find justification however when they contribute to the best possible distributions. This Chapter explored the guiding value of two alternative metrics of distributive equality: equality of primary goods and equality of capabilities.

EQUALITY OF PRIMARY GOODS – The primary goods (basic liberties, opportunities, wealth and income; and social bases of self-respect) are what free and equal persons need as citizens in order to live a complete life.[1220] These goods are necessary for people to develop and exercise their two moral powers (a capacity for a sense of justice and a capacity for a conception of the good). They are also a type of all-purpose means that enable persons to freely form, revise, change and pursue of a (sufficiently) wide range of personal ends. The primary goods look to respect each person's

[1218] Reuben Binns et al., '"It's Reducing a Human Being to a Percentage"; Perceptions of Justice in Algorithmic Decisions', *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems - CHI '18*, 2018, 5, https://doi.org/10.1145/3173574.3173951; Lee, 'Understanding Perception of Algorithmic Decisions'.

[1219] Cohen, 'On the Currency of Egalitarian Justice'; Schemmel, 'Distributive and Relational Equality'; Voigt, 'Distributive Equality, Relational Equality and Preferences about Higher Education'.

[1220] Rawls, *Justice as Fairness: A Restatement*, 58.

inviolable capacity towards freedom, self-development, and self-determination. The distribution of primary goods is governed by two principles of justice, which try to offset the impact morally arbitrary contingencies have on people's life's prospects (social class of origin, natural talents, and abilities, and good or bad fortune)

Algorithmic structures can be evaluated in light of their contribution to the desired distribution of primary goods as defined by the principles of justice. First, *every* person has an indefeasible claim to access a set of equal, essential liberties that are vital for the development and exercise of their fundamental interests. The actions of decision-makers, including the differentiation they impose and the socio-relational and socio-economic inequalities they produce, should not interfere with people's access to these liberties, nor limit certain persons or groups in their equal enjoyment of these liberties.

Second, the algorithmic environment should be structured and regulated so that fair equality of opportunity can be maintained. Within the algorithmic environment, and the various decisions that take place therein, historical, and social fortune should not negatively affect people's prospects. Instead, under fair equality of opportunity, people with the same abilities, talents, willingness, and motivation, should have an equal chance of success. It was demonstrated however that the application of fair equality of opportunity is not always as straightforward. For instance, environmental conditions and social attitudes affect a person's capacity to develop their talents and abilities. In order to ensure fair equality of opportunity, and assess and overturn the influence of social class, algorithmic decision-makers should consider how the talents, abilities, and motivations they have identified as relevant for a given position or opportunity, are shaped, and influenced by people's social conditions. Furthermore, as fair equality of opportunity is motivated by the ideal of equal social standing, particular attention must be paid to socio-relational egalitarian harms, and associated vulnerabilities, algorithmic structures and decision-making acts risk imposing onto decision-making subjects.

Third, in order to mitigate the morally arbitrary impact natural talents, abilities and fortune have on a person's social position, the socio-relational and socio-economic inequalities produced by algorithmic structures and decision-making acts should be to the benefit of those who are least-advantaged. During the design of a system, decision-makers should prioritise those choices and options that have the capacity to make the least-advantaged best off. In order to tackle the various socio-relational and socio-economic harms different systems can impose onto diverse groups, the notion of the least-advantaged should be approached with some degree of flexibility.

Finally, it was argued that algorithmic acts of oppression and domination, including certain algorithmic classification acts, pose several identity-related risks that can undermine a person's self-confidence and self-worth, and therefore unjustifiably interfere with people's access to the social bases of self-respect.

EQUALITY OF CAPABILITIES – Following a capability approach towards justice, differences between people can be evaluated in terms of the real opportunities people have to choose between and achieve (combined capabilities) alternative beings and doings (functionings). These real opportunities are created by a combination of a person's and the political, social, economic, and socio-technical environment in which they reside. In order to lead a dignified life of truly human functioning, Nussbaum believes all people should have access to a set of central or core capabilities (life; bodily health and integrity; senses, imagination and thought, emotions; practical reason; affiliation; other species; play and control).

Having become an integral component of people's social world, algorithms are best considered an environmental factor that can either contribute to, or limit, people in their substantial freedoms to achieve valuable functionings. For instance, the capabilities approach affirms the right to self-determination and self-development as fundamental interests of people: people should not only have the freedom to pursue their own ends in accordance with their choice, but also be given the suitable external conditions so they can transform and exercise their skills in socially meaningful and recognised ways. In this context, algorithmic acts of domination and oppression function as an external constraint that denies people the opportunity to flourish.

Against this background, the central list can function as an evaluative benchmark to evaluate the impact algorithmic decision-making structures and acts, and the (in)equalities they produce, have on people's substantial freedoms. The central capabilities urge decision-makers (and regulators) to be outward looking, beyond insular decision-making contexts. Where a decision affects the enjoyment of multiple central capabilities, the exclusionary harms they risk imposing, among others through the socio-relational and socio-economic inequalities they generate, tend to be greater. Considering the unique nature of the algorithmic environment, they should identify whether over time the outcomes they produce can be corrosive or fertile to people's overall freedom and choice.

The algorithmic environment should be structured so it can accommodate the dynamicity and diversity of human needs. Algorithms could be designed with a goal of dismantling and correcting the deep social injustice, structural inequalities or otherwise harmful disadvantage, certain

persons, and groups face. In this context, inclusive and (democratic) participatory design strategies were identified as one potential approach towards equality of capability that require decision-makers to understand and learn from the lived experiences of their decision-subjects.

Algorithmic decisions, which often rely on group-level of reasoning, might lack the required specificity however to address the diverse nature of human needs. Whereas they could provide an initial basis for decision-making, they might need to be combined with mechanisms that enable decision-subjects to have their particular needs recognised, heard and if necessary, accommodated. Only then could the social goal of equality of capability *for each person* be achieved.

PART IV- Reconsidering the Concept of Equality for the Evaluation and Regulation of Algorithmically Guided Decision-Making: Concluding Remarks and Recommendations

8 Repositioning the Concept of Equality for the Regulation of Algorithmically Guided Decision-Making: Concluding Remarks and Recommendations

In this concluding Chapter, I consolidate my main research findings. I first provide a brief overview of the research context, question, and methodology (Section 8.1). Following the structure of this dissertation, I then demonstrate how the concept of equality was operationalised for the evaluation and regulation of algorithmically guided decision-making systems (Section 8.2). Based upon my findings, I formulate a set of normative recommendations to reposition the concept of equality within the algorithmic governance debate in an effort to further strengthen its guiding function for the evaluation and regulation of algorithmically informed decision-making systems (Section 8.2.4.).

8.1 Research Context, Question and Methodology[1221]

RESEARCH CONTEXT - Algorithms are increasingly relied upon to help decision-makers automate, streamline, structure and guide a variety of decision-making processes in both the public and private sector. As algorithms have become deeply entwined with our everyday living and social environment, it is only reasonable to question whether their introduction in knowledge and data-driven decision-making purposes may change people's being in the world in novel ways. This dissertation has been concerned with one particular type of injustice that may arise out of the deployment of algorithmic decision-making system: the emergence of unjust (in)equality brought about by differentiation acts that take place within and as part of these systems. Due to their complexity however, it has become increasingly difficult to assess whether, and under what conditions, algorithmically guided decisions, and the inequalities they generate, interfere in our lives in ways that are fair or unfair, justified, or unjustified.

[1221] For a thorough overview of the research context, question and metholody, including research choices made, I refer to this dissertation's introductory Chapter "Reconsidering the Concept of Equality?"

RESEARCH QUESTION, OBJECTIVE AND STRUCTURE - Algorithms present an articulation challenge which limits our ability to adequately evaluate and regulate their functioning. In this context, it was hypothesised that a renewed investigation into the concept of equality, genuinely informed by the (societal) challenges algorithms risk to introduce, may improve both our understanding of algorithmic injustice and the regulatory response this may prompt. This led to the formulation of the following research question: (How) can we operationalise the concept of equality to guide the evaluation and regulation of algorithmically guided decision-making processes in light of the inequalities they produce?

To achieve the main researching objective of revitalising the concept of equality, I developed a multidimensional understanding of equality, which was operationalised as a practicable and interpretative guiding mechanism by way of performing three particular functions, namely:

a) To identify and articulate when, and under what conditions, algorithmically guided decision-making processes, the instances of algorithmic differentiation that take place therein, and the inequalities that flow thereout, risk being fair or unfair, justified, or unjustified **(identification).**

b) To assess and evaluate specific interpretations of equality and their response to the risk of egalitarian harm algorithms present **(evaluation).**

c) To deepen our understanding of the egalitarian harms algorithms risk to impose, the conditions under which these harms manifest in practice, and the institutional safeguards needed to address them **(navigation).**

In order to achieve the main research objective, and demonstrate how equality can be operationalised, I first examined the (practical) functioning of algorithmic decision-making systems and the particular societal challenges these systems risk to introduce **(Chapter 1)**. Drawing inspiration from both European equality and non-discrimination law and political philosophical theories of justice, I then developed my understanding of equality as a concept that can be further specified alongside three interrelated axes: a moral, socio-relational, and distributive dimension **(Chapter 2).** In a first step, I positioned the multidimensional model against the algorithmic environment in order to articulate and identify various egalitarian harms algorithms risk to impose **(Part I: Identification, Chapter 2)**. In a second step, the multidimensional model was used as a supportive mechanism to unearth the legal conceptualisation of equality found within European equality and non-discrimination law. By

concretising whom the law protects against which (algorithmic) inequalities in the pursuit of which ideals, I then assessed the law's approach towards equality in terms of its ability to address the algorithmic injustice risk to introduce **(Part II: Evaluation; Chapters 3-5).** Finally, the model was relied upon to locate specific notions of equality chosen due to their correspondence with the aforementioned dimensions. These socio-relational (domination and oppression) and distributive (primary goods and capabilities) notions were examined in order to concretise the egalitarian harms algorithms risk to produce, the conditions under which these harms may manifest in practice, and the safeguards that can be provided to protect decision-subjects and society at large against them **(Part III: Navigation; Chapter 6 and 7)**. In this final Part, I summarise the main research findings, and derive from them, a set of normative recommendations to bolster the concept of equality's position as a guiding notion in the algorithmic governance debate **(Part IV: Synthesis; Chapter 8)**.

8.2 Main Research Findings

8.2.1 Part I - (In)equality in the Algorithmic Environment: Identifying The Risk of Egalitarian Harm.

The Algorithmic Environment

In their function as a decision-making support mechanism, algorithmic techniques are often used for two broad purposes: knowledge creation and/or knowledge application. Assisted by novel analytical techniques, decision-makers try to uncover similarities between persons and situations to motivate their like treatment, as well as the differences between persons and situations that can justify differential treatment. For decision-subjects, these differentiation acts may generate significant consequences regarding their social position and life's prospects. When applied on a large enough scale, they may also initiate significant societal change. Non-automated, human decision-making processes may affect individuals and initiate social change too, however. If we want to successfully operationalise the concept of equality for the evaluation and regulation of algorithmically guided decision-making systems, we must clarify the distinctive nature of algorithmic (in)equality. In other words, we must understand and recognise the particular challenges algorithms may introduce and the environment in which they operate.

By examining their technical functioning, and the advantages decision-makers derive from them, I identified a series of properties algorithmic decision-making techniques exhibit that help explain

the distinctive nature of the inequalities they produce. First, in their function as knowledge generators, the decisional grounds they uncover can be novel, counterintuitive, and seemingly random. Second, due to the computing power at their disposal, algorithmic decisional grounds can take on more complex forms. Third, knowledge is derived from and/or applied onto group structures. Though an individual decision-subject may be the final recipient of a decisional outcome, that outcome is often based upon a generalised assumption or profile constructed from the behaviours, characteristics, or preferences of others with whom the decision-subject shares an alleged similarity. The decisional outcome can therefore be uniformly applied onto every decision-subject who fits the profile, and when automated, this process is easily streamlined. Fourth, these technologies can be rolled out on a large-scale basis. In addition, due to the algorithmic environment's interconnected nature, the consequences of algorithmic (in)equality may take effect in multiple areas of life, and as a result thereof, intensify over time. These two structural characteristics further exacerbate the risks of scale already associated with the use of profiles. By virtue of these four properties, algorithmic inequalities, and the social change they initiate, may differ in both content and form: they can restructure society alongside new and existing social dimensions. Finally, control over this restructuring capacity may become centralised in the hands of those who hold power over the means of knowledge discovery, production, interpretation, and application. Their choices significantly impact how society will change.

It would be wrong however to call algorithmically guided decisions-making systems purely passive artefacts during this process. While they incorporate meaning, they also shape meaning. Algorithms take meaning from their environment in the sense that they take data and learn from the outside world, follow the instructions of their designers, adapt to their environment through their interaction with end-users and decision-subjects, etc. At the same time, once deployed, their impact may be transformative, unforeseen, and unexpected. First, ample evidence indicates how algorithms not only replicate, but also amplify and reinforce the historical and social injustice faced by social groups. Second, and as hinted at before, when decision-makers rely upon novel or hitherto undiscovered decisional grounds, they have the potential to restructure society alongside less tangible social dimensions and expose decision-subjects and society at large to new forms of injustice that are non-representational. The injustice strikes us as unfamiliar, and hence more difficult to articulate, when it appears in a form and shape we have not yet witnessed or experienced in the past.

Equality as a Multidimensional Concept

EQUALITY AS A MULTIDIMENSIONAL CONCEPT - After having defined the algorithmic research environment, I developed an understanding of the concept of equality that could be operationalised to make sense of the injustice we find therein. A society's commitment to equality reflects its intent to equalise the enjoyment of certain fundamental interests for all people. Informed by political philosophy and legal theory, I argue that equality is best approached as an *abstract, normative* concept that can be further *interpreted* or *conceptualised* alongside three interrelated axes: a moral, socio-relational, and distributive dimension. Each dimension is associated with a particular (policy) ideal we can ascribe to equality as a social value that informs the desired structure of society. Through their interpretation of these related, but nonetheless distinct dimensions, a law, policy, or theory can give normative substance to the concept of equality. These specific conceptualisations of equality will concretise what fundamental interests ought to be equalised within society and how this equalisation process can be achieved. Likewise, they specify which inequalities are unjust.

THE MORAL DIMENSION - As a moral ideal, equality reflects a society's commitment to affirm every person's equal moral worth: people should be treated with equal concern and respect. The moral dimension is foundational in the sense that any given framework, law, or policy can either accept or denounce the ideal of moral equality. Once affirmed, a policy, framework or theory can give expression to the ideal that people are one another's equals by giving normative substance to the socio-relational and/or distributive dimension. In this regard, a commitment to equality as a moral ideal legitimises an inquiry into the social, relational, or material differences that exist and arise between people.

THE SOCIAL DIMENSION - The ideal of social equality reflects the belief that every person must be respected, recognised, and valued as a free and equal member of society. The socio-relational dimension therefore concerns a society's willingness to ensure people's fundamental interest to be recognised as equals in the various relationships they hold. On the one hand, the socio-relational dimension pertains to the (positive) institutions, principles, conditions, practices and rights defined within a society to ensure people can maintain their status as social equals and stand in equal social relations.[1222] On the other hand, this dimension concerns a society's attempt to abolish (negative) actions, treatment, behaviours, structures and phenomena that enable

[1222] Anderson, 'What Is the Point of Equality?', 313; Scheffler, 'The Practice of Equality', 21–22.

individuals, (social) groups, private corporations or governments to threaten or undermine people's status as social equals.

THE DISTRIBUTIVE DIMENSION – When equality is positioned as a distributive ideal, a just society will be structured with the purpose of maintaining a desired distributive state of affairs. People should have (roughly) equal access to, or the enjoyment of, certain justice-relevant goods. These justice-relevant goods represent a certain fundamental interest that people share and any interference in their distribution is therefore an interference, either negative or positive, with that interest. Importantly however, following a distributive paradigm, we can assess the justifiability of (in)equality between people by examining whether the advantage and disadvantage linked to them contributes to the desired distribution of justice-relevant goods. In this regard, the distributive dimension is concretised by specifying a series of justice-relevant goods that are deemed critical for people and formulating (institutional) rules regarding their preferred allocation.

Identifying Egalitarian Harm

Before evaluating algorithmic decision-making systems, and the regulation thereof, I first had to identify the algorithmic egalitarian harms that threaten decision-subjects and society at large. To that end, the multidimensional model was positioned against the algorithmic environment. Because the model represents a core set of social ideals commonly associated with the notion of equality as a structuring value, it could be used to evaluate the restructuring capacity of algorithms, and in particular, the egalitarian harms these restructuring activities risk to impose. Every time an algorithmic decision interferes with one of the identified dimensions of equality, they risk interfering with the desired structuring of society. Their capacity to do so was identified as a risk of egalitarian harm. The (un)fairness and (in)justice of algorithmic structures and decision-making acts must be located, articulated, and evaluated in terms of the harms algorithms are likely to produce. The regulatory response specific interpretations of equality provide against algorithmic injustice can be evaluated in light of their ability to address the identified risk

THE RISK OF MORAL HARM - Algorithmic decision-making acts and structures generate a moral egalitarian harm when they do not treat decision-subjects with equal concern and respect. In practice, whether an algorithmic decision fails to do so hinges upon our specific (socio-relational and/or distributive) interpretation of equality. Still, by affirming our commitment to moral equality we legitimise our inquiry into algorithmic decision-making: we can review any

algorithmic act and (in)equality in order to verify whether they respect decision-subjects as beings of equal moral worth.

THE RISK OF SOCIAL HARM – Algorithms impose social or socio-relational egalitarian harms when they place decision-subjects in relationships of unequal social standing or social inequality. The risk for decision-subjects to be exposed to socio-relational egalitarian harm may arise on two fronts. First, decision-subjects often have little control over the means of knowledge discovery, production, interpretation, and application. Without appropriate safeguards, the risk exists that decision-makers take unfair advantage of their position of power, and that decision-subjects are unable to challenge the control decision-makers exercise on equal footing. Second, algorithmic decision-makers co-mediate the relationships decision-subjects hold vis-à-vis themselves and others. Decision-makers have the power to divide the population into broad consequential categories. If decision-subjects do not have any institutional support, decision-makers are in full control over the conditions under which this social division occurs. Depending on the treatment imposed, and/or the inequalities produced, a decision-subject or algorithmic group can see their social status lowered, maintained, or increased.

THE DISTRIBUTIONAL DIMENSION – As algorithms co-mediate and co-shape society's basic institutions, they also co-govern the distribution of justice-relevant goods. Algorithmic decision-making acts and structures, among others through the socio-relational and socio-economic (in)equalities they produce, impose distributive or allocative egalitarian harms when they interfere with the desired distribution thereof. Conversely, algorithmic actions are justified when they promote just distributions.

EQUALITY OF WHOM? – Considering the unique properties of the algorithmic environment and given the wide variety of domains in which algorithms can find application, even the use of seemingly random and non-representational differentiation grounds can generate substantial differences in decision-subjects' enjoyment of their fundamental socio-relational or distributive egalitarian interests. Put simply, egalitarian harms are risked, both by representational social groups and non-representational algorithmic aggregates.

Of course, most societies accept and tolerate the fact that differences between people can and will exist. Put differently, they must define when and under which conditions those differences can be justified. Still, if there are certain fundamental interests that a just society should guarantee to each and every person, the protection offered against the incurrence of algorithmically produced egalitarian harms should not be limited to social groups and their group members, but also extend

to algorithmic aggregates and their group members. Doing otherwise would gravely underestimate the capacity algorithms have to restructure people's relationships and their access to justice-relevant goods, and the egalitarian harms decision-subjects risk to suffer as a result of these operations. Hence, a robust and future-proof approach towards equality should be able to address both types of groups in their disadvantage.

8.2.2 Part II - Evaluating the Law: The Legal Conceptualisation of Equality and the Risk of Algorithmic Egalitarian Harm.

Considering the restructuring capacity of algorithms, and the egalitarian harms algorithms risk imposing on decision-subjects, in this dissertation's second part, I examined the legal conceptualisation of equality found within the Council of Europe's and European Union's legal orders. The research objective was twofold: to uncover and subsequently evaluate the legal conceptualisation of equality by positioning it against the algorithmic environment and the risk of egalitarian harm previously identified.

A METHODOLOGY – To identify the strengths and weaknesses of equality as a legal concept for monitoring the algorithmic environment two lines of inquiry were researched. Because the law tends to define equality as the absence of discrimination, it can first be examined whether a legal system has placed definitional constraints onto the notion of discrimination. When legal systems follow an open model approach, any differentiation act can constitute an unjustified violation of equality. When a closed model approach is adopted, not every act of differentiation can be made subject to review, but only those who meet certain definitional conditions. These conditions function as a filtering mechanism. As algorithmic differentiation grounds represent a totality of decision-subjects, I argued that such restrictions effectively limit who, and *in casu*, what algorithmic groups (and their group members) can launch an equality claim. Second, once it has been determined that a differentiation act can be made subject to legal scrutiny, the evaluation mechanisms present within the law, as well as the (normative) components that are taken into account during this justification process, will further determine whether, and to what extent, the law can, and is willing to address, algorithmic egalitarian harms.

These axes of inquiry were inspired by the multidimensional concept of equality. Through their investigation, one can uncover whom the law protects (equality of whom?) against which inequalities (what egalitarian ideals do they promote?) and how (what mechanisms have been

put when equality is positioned against competing values?). Furthermore, by focusing on these areas of inquiry, it was possible to isolate and evaluate those (normative) choices made by European lawmakers and the judiciary that affect the law's ability to address algorithmic injustice.

GENERAL CONTEXT - Within the European legal order, the principle of equality performs a variety of functions: since its beginnings as an infrastructural or institutional check against inequalities resulting from arbitrary or random public decision-making, the principle of equality has gained a more transformative function aimed towards the realisation of specific socio-relational and distributional egalitarian goals. In both orders, equality shares a formal foundation that has complemented with a more substantive vision on equality that is sensitive to the social context in which disadvantage, discrimination and inequality can arise. As a notion subject to social change, equality is often caught between various (competing) narratives and interests. Consequently, the legal notion of equality and non-discrimination rules have not been always applied in a clear, consistent, or coherent manner. Despite this complexity, and if the legal conceptualisation of equality is to become a guiding principle for the future regulation of algorithmically guided decision-making, and the interpretation of algorithmic egalitarian harm, some conclusions could be drawn regarding its ability to do so.

THE DEFINITION OF EQUALITY AND NON-DISCRIMINATION – In their definition of equality and non-discrimination law's personal scope, both the CoE and EU system maintain a grounds-based approach. As a preliminary observation, the argument was made that these approaches - which rely upon the identification of the attributes or characteristics used as a basis for differentiation - face difficulties when positioned against the algorithmic environment. First, algorithmic decision-making processes operate in covert and dynamic ways. The attributes used to differentiate between individuals or groups of individuals at a given point in time might not always be known. Second, in an environment where novel differentiation grounds can be generated on a continuous basis and applied in a vast variety of domains, the differentiation grounds themselves might hold little (initial) evaluative value. Instead, during the evaluation of algorithmic structures and decision-making acts extra attention should be paid to the particular egalitarian harms systems (risk to) impose and how these harms may conflict with the core ideals pursued by an equality-led society.

While a grounds-based approach may constrain the law's ability to address algorithmic harm, additional definitional limitations have been placed onto the notion of discrimination. The ECtHR does not prohibit all forms of differential treatment, but only those that are based upon an objective, identifiable or personal characteristic or "status". The notion status must be given "wide

meaning" in the sense that it does not only refer to personal characteristics that are innate or inherent. The notion of discrimination is therefore quite broad and flexible. Still, it remains unclear whether algorithmic differentiation grounds fit within those conceptual boundaries. Opaque, dynamic, (seemingly) random, intangible, or non-representational profiles for instance appear neither identifiable, nor personal. In the EU, the CJEU has shown a general reluctance to extend the (interpersonal) protection offered to grounds the EU was not (expressly) mandated to govern - though equality does appear to have a more open-ended and sweeping scope where it functions as an institutional check. Furthermore, and despite having shown awareness for the disadvantage faced by persons that are multiply burdened, in both systems, the understanding of more complex forms of discrimination, including intersectional discrimination, remains underdeveloped. Hence, one has to wonder whether the inequalities produced by algorithmic decisions, when the differentiation grounds in question are multidimensional and complex in nature, could be recognised as discrimination. In sum, as the law tends to classify as discrimination those differentiation acts that are based upon identifiable or personal grounds that are (likely to be) recognised as having a certain degree of social significance or social saliency, harms faced by non-representational algorithmic aggregates, whether they are defined by complex profiles or not, are unlikely to find protection following a European-inspired definition of discrimination.

JUSTIFICATION OF EQUALITY AND NON-DISCRIMINATION GROUNDS – Concerning the evaluation and justification of differential treatment, both systems depart from a classic and (formal) formulation of the principle of equality: the differential treatment of comparable situations (or vice versa, similar treatment of incomparable situations) can be justified as far as a reasonable and objective justification can be provided. Under this formulation, which generally entails an investigation into the legitimacy of the objectives pursued, and an assessment of a measure's proportionality, the Courts engage in a balancing of competing interests, the outcome of which determines whether an act of discrimination, and the resulting inequality, constitute an injustice.

As part of this exercise, a stricter level of scrutiny is typically applied where differential treatment is based upon historically or socially significant suspect or protected characteristics. In these instances, decision-makers must advance stronger reasons to justify differential treatment. Across both systems, a similar pluralist approach was identified regarding the underlying rationales that motivate this heightened level of review.

Certain motivations tend to be more strongly associated with a formal understanding of equality. For example, differentiation should be treated with a higher degree of caution when it is based on

irrelevant or immutable characteristics. Such rationales fail to articulate algorithmic egalitarian harms for two reasons. First, from a conceptual perspective, it is difficult to define the circumstances under which (non-representational) algorithmic differentiation grounds are irrelevant or immutable. Second, these rationales tend to locate the wrongful character of discrimination in the nature of the characteristic upon which differentiation was based: a characteristic is either relevant or irrelevant, mutable, or immutable, and depending upon the answer, its use can be justified or unjustified. Such approaches risk being agnostic to the particular social conditions under which (algorithmic) egalitarian disadvantage comes about. Hence, they might overlook the particular challenges faced both by historically disadvantaged social groups and non-representational algorithmic aggregates. By contrast, when substantive rationales are invoked to ground equality, the law does try to capture and account for contextual and social indicators of inequality. The wrongful nature of discrimination is not tied to the nature of the differentiating trait. Certain persons and groups face prejudice, stereotypes, and stigma, or they find themselves in a vulnerable position. Differentiation is forbidden when the use of a given characteristic is reflective, or likely to perpetuate or reinforce this disadvantage. These rationales are tied to (social) groups who have faced a history of disadvantage. Such a history of disadvantage is not something non-representational algorithmic aggregates are likely to have experienced. Hence, it is unclear whether they too could benefit from more contextual and disadvantage-oriented reasoning.

Under both formal and substantive rationales, equality has come to represent a social value that represents the lawmaker's commitment to guard people against unjustified (social) exclusion. In differing ways, equality is put forward as a mechanism towards the securitisation of people's equal social standing, as well as their equal access to certain justice-relevant goods. In addition, the principle of equality has also been linked to other values, such as the respect for (individual) human dignity and autonomy. Admittedly however, the substantive rationales show a more clear and deep concern for the socio-relational (misrepresentation and lack of recognition) and distributional (material deprivation) harms certain social groups face. Yet, following a grounds-based, suspect-centric approach however, the language of European equality and non-discrimination laws - including the way (social) disadvantage has been framed through the above-mentioned rationales – has been constructed towards the protection of social or socially salient groups, rather than non-representational groups or aggregates.[1223]

[1223] See on this point: Wachter, 'The Theory of Artificial Immutability'.

Moreover, when discrimination claims pertain to decisions that directly affect people's enjoyment of specific prized goods, such as human rights, education or employment, socio-relational inequality may not always be recognised as a stand-alone harm both within and outside these distributive contexts. Furthermore, due to the interconnected nature of the algorithmic environment, a decision-making system may introduce inequality in one area of life, whereas its actual impact takes place in another. By evaluating single decision-making instances, the law may overlook how people's social position and their enjoyment of certain justice-relevant goods is determined by a totality of interweaving relational and economic actions, behaviours, and practices. To adequately map these dynamics, inequality must be evaluated within a broader context.

Whether the current approach towards equality will be maintained in light of algorithms' transformative impact remains to be seen: due to both system's evolutive approach, the interpretation of equality, and its conceptual boundaries, are subject to change. For instance, although unexplored in practice, the equality clauses enshrined within the ECHR and the CFR are open-ended. This could allow the ECtHR and CJEU, within the confines of their respective competences, to evaluate the more novel egalitarian harms produced within the algorithmic environment. Certain structural elements of the classic (and formal) equality review do not befit the algorithmic environment, however. For instance, equality is often formulated as a principle that mandates a comparison of situations, but this approach faces obvious challenges. It is often difficult to apply a comparative analysis onto more complex forms of social disadvantage, such as intersectional discrimination. In addition, the comparative review runs counter to the functioning of algorithms: algorithms often have as their exact goal to place people in comparable and non-comparable situations. Furthermore, certain egalitarian harms decision-subjects may experience, such as their exposure to negative attitudes as a result of untruthful generalisations, stand on their own, whether other decision-subjects have gained a comparative advantage or not.

Yet, even if the Courts are in the possibility to review algorithmic inequalities, valid concerns have been raised regarding their legitimacy to do so. First, given the novelty of the challenges involved, fundamental questions regarding the structuring of society would be moved away from public discourse and placed into the hands of judicial authorities instead. The risk of judicial activism would be lower if the courts were to follow a well-informed, clearly structured, and consistently applied proportionality test. Second, question marks could still be raised as to the desirability to alter the current approach of (substantive) equality and non-discrimination laws. Though arguably flawed, a system constructed around the definition and identification of protected characteristics has benefits. The protected grounds offer legal certainty and clarity, and through

their presence, non-discrimination laws gain an expressive and symbolic value. Furthermore, once specific acts of differentiation have been clearly identified as unjustified or morally reprehensible, resources towards the enforcement of equality could be allocated more efficiently to benefit those social groups who need it most.

If the current legal conceptualisation of equality were to be positioned as a bulwark against algorithmic inequality, it would be insufficiently robust to address the egalitarian harms algorithmic decision-making structures and systems risk imposing onto (members) of more intricately defined socially disadvantaged groups or non-representational algorithmic aggregates. The core rationales that have been used to articulate who will be protected by equality and non-discrimination laws and why (irrelevance, immutability, historic prejudice, etc.), might make sense in relation to social groups, but not in relation to non-representational ones.[1224] Yet, the harms both groups experience could very well be the same.

European equality and non-discrimination laws are therefore at a crossroads regarding the function of equality in the future regulation of algorithmic differentiation. More specifically, it must be assessed whether the grammar and vocabulary of European discrimination laws should be altered in order to accommodate the socio-relational and distributional egalitarian harms non-representational groups may face. To that end various options are open. An open-ended equality clause could be introduced, whereby the review process can evaluate decision-making systems on the basis of the disadvantage they (are likely to) produce.[1225] When doing so, it could also be examined to what extent the harms faced by non-representational groups can be clarified by, and incorporated into, the existing language of equality and non-discrimination law. In addition, the material scope of the law could be expanded as to better accommodate the interconnected nature of the algorithmic environment. Or instead, a *sui generis* framework could be considered in order to ensure that people's enjoyment of certain fundamental interests can be equalised. When making these reflections, we can turn to political philosophical notions of equality to further strengthen our approach towards equality.

[1224] Wachter.

[1225] See also: Gerards and Zuiderveen Borgesius, 'Protected Grounds and the System of Non-Discrimination Law in the Context of Algorithmic Decision-Making and Artificial Intelligence', June 2022.

8.2.3 Part III - Navigating Political Philosophy: Socio-Relational and Distributive Conceptualisations of Equality and the Articulation of Algorithmic Egalitarian Harm.

In Part III, I located specific political philosophical conceptualisations of equality chosen due to their correspondence with the dimensional framework. By transposing various socio-relational (domination and oppression) and distributive (equality of primary goods and capabilities) interpretations of equality onto the algorithmic environment, I was able to elaborate on a series of egalitarian harms algorithms risk to impose, and how these can arise in practice.
These notions specify how equality can be operationalised as a socio-relational and/or distributive ideal within the algorithmic environment. Under the assumption that the social restructuring capacity of algorithms should be of utmost concern, these notions are an appropriate starting point to reflect upon a set of alternative or additional regulatory action points. Moreover, as European equality and non-discrimination laws too aim to redress socio-relational and distributive inequality, the identified notions can complement and strengthen the legal approach in a more foundational, abstract, and principled way. They can offer lawmakers and the judiciary additional anchoring points to ground the normative foundations of equality law. At the very least, they give further insight into the potential egalitarian harms algorithms risk to impose, the conditions under which these risks may manifest, and the institutional safeguards they could put in place to prevent or address them.

Having identified the deficiencies of the current legal framework, these notions of equality have a number of immediate advantages. Their protective scope is not limited by, nor based upon, a defined list of (identifiable) characteristics or traits. Instead, their protective scope is defined by the underlying fundamental interest that needs to be equalised. If every person has an interest to be either treated with equal concern and respect; valued and recognised as being of equal social standing; and/or given (roughly) equal access to certain justice-relevant goods, any algorithmic operation that is liable to interfere with these interests represents a risk of harm to the decision-subjects involved. Regardless of the decisional ground relied upon and the groups that were captured, as soon as these interests are interfered with, decision-subjects should be given the possibility to launch an equality claim. Furthermore, these interests matter outside the selectively identified decision-making domains one typically associates with substantive non-discrimination laws. Hence, decision-subjects should be able to contest unjust (in)equality irrespective of the domain, area, or sector where the (initial) interference took place or effect. In other words, range

of protection is therefore more likely to transcend the boundaries of equality and non-discrimination laws.

SOCIO-RELATIONAL EQUALITY AND ALGORITHMS – Drawing on the work of Pettit and Young, I first examined the socio-relational notions of oppression and domination. Each person has an equal entitlement to lead their life free from oppression and domination. Acts of oppression and domination constrain persons in their self-development (capacity to learn, develop, exercise, and express skills, capacities, and experiences in socially meaningful and recognised ways) and self-determination (the capacity to exercise control over, and participate in determining, their own options, choices, and the conditions of their actions) respectively. In order to realise social equality within the algorithmic environment, decision-subjects should be guarded from structures and acts of algorithmic oppression and domination. At the same time, decision-subjects must be resourced so they can meaningfully exercise their right to self-development and self-determination.

Decision-subjects, who are often powerless in their confrontation with decision-makers, are subject to algorithmic oppression when they are a) exploited, b) excluded from, or over-exposed to, a decision-making process (marginalisation), c) rendered invisible due to the capacity of decision-makers to universalise their world views as the norm as result of their access to the means of interpretation and communication in the algorithmic environment (cultural imperialism), or d) exposed to a digital replication of the systemic forms of violence they already experience (violence). Decision-subjects are subject to algorithmic domination when the decision-maker has the capacity to interfere into their valuable options and choices in a manner that is arbitrary, unchecked, or uncontrolled. In addition, I argued that algorithmic generalisation acts, when they reflect a (negative) stereotype pertaining to a specific person or group, or more broadly, when they are the sole means decision-makers rely upon to predict, treat, and judge decision-subjects on, can actively undermine a person or group's social standing. Incontestable generalisations constrain, deindividualise, or misrepresent the decision-subjects they capture through pre-determined labels. Because decision-subjects cannot exercise control over the conditions that govern their actions and behaviours, nor express their point of view and experiences, they are denied self-authorship. This too can be experienced as oppressive and/or dominating.

Oppression and domination may arise either directly or indirectly. First, through their position of power, decision-makers may directly oppress or dominate decision-subjects. Their ability to do so signals an inegalitarian social relationship. In addition, these situations are indicative of a

regulatory environment where decision-subjects have been given insufficient resources to withstand the abusive exercise of power by decision-makers. Second, as algorithmic structures and decision-making acts have become integrally linked to society's basic institutions, they can actively restructure the relationships people stand in, among others when their deployment co-governs the burdens, benefits and opportunities decision-subjects are afforded. Decision-makers should be aware that as a result of these actions, they can increase, maintain, or decrease a person or group's social standing. Due to this relational impact, they may indirectly contribute to the oppression and domination of certain groups within society. Under both circumstances, decision-subjects may find that their behaviour and actions are (unjustifiably) curtailed in a manner that disallows them to participate to their social environment as equals: they are rendered voiceless, invisible, uncertain, cautious, and less autonomous. They are placed in a position of heightened vulnerability. Or put simply, they suffer a status-inequality.

Of course, due to society's increasing reliance on data-driven technologies, algorithmic decision-makers have been placed in a position where they can easily interfere with the way people exercise their right to self-determination and self-development. Does this mean their deployment should be abolished? Not at all. From the negative framing of socio-relational equality as an obligation of decision-makers to avoid the imposition of status-harms, conclusions were drawn as to the positive institutions, principles, conditions, practices and rights people need to enjoy, or have access to, in order to equalise power relationships within the algorithmic environment so that people can adequately exercise their right to self-development and self-determination.

DISTRIBUTIVE EQUALITY AND ALGORITHMS – Though they often share a strong concern for deeply rooted and entrenched socio-relational and institutional inequalities, rather than directly focus on the relationships people stand in, distributive egalitarians maintain equality through the identification of a set of (non-)material justice-relevant goods the access to which should be guaranteed on a (roughly) equal level. Those justice-relevant goods are needed to secure a given fundamental interest people are said to share. The outcomes people end up with are the focal point of this evaluation. The outcomes algorithmic structures and acts produce can thusly be evaluated in terms of their impact on people's holding of justice-relevant goods. Algorithmic decision-making structures and acts, including the socio-relational and socio-economic inequalities they generate, will produce distributive harms when they fail to contribute to the desired distribution. In this regard, I examined two distributive conceptualisations of equality and

how their application would affect the evaluation and regulation of algorithmic decision-making practices: equality of primary goods and equality of capability.[1226]

EQUALITY OF PRIMARY GOODS - UNDER a Rawlsian view, free and equal people need access to a set of primary goods (basic liberties, opportunities, wealth and income, and the social bases of self-respect). On the one hand, these justice-relevant goods enable persons to develop and exercise their two moral powers (a capacity for a sense of justice and a capacity for a conception of the good). On the other hand, they give persons the ability to freely form, revise, change and pursue a broad range of personal ends. At the same time, the algorithmic environment should be structured so that the impact of morally arbitrary contingencies - social class of origin, natural talents, and abilities, and good or bad fortune - on people's life prospects is counterbalanced. The latter can be achieved by making sure that the outcomes algorithmic structures and decision-making acts produce are in accordance with the principles of justice. These define the desired distribution of justice-relevant goods.

According to the first principle of justice, each person has an indefeasible claim to a set of essential equal basic rights and liberties. Algorithmic structures and decision-making acts should not (unjustifiably) interfere with people's enjoyment of these liberties, nor, through the inequalities they produce, deprive, or reduce certain persons or groups in their access to them.[1227] Second, social and economic inequalities ought to be attached to positions open to all under conditions of fair equality of opportunity. Both algorithmically mediated structures, as well as isolated decision-making acts, should not let historical and social fortune negatively affect people's life prospects. Instead, people with the same abilities, talents, willingness, and motivation, should have an equal chance of culture and achievement. Finally, to correct the arbitrary influence that natural talents, abilities, and fortune have on people's social positions and opportunities, algorithmically generated socio-relational and socio-economic inequalities should benefit the least advantaged members of society. In this context, the regulation of the algorithmic environment, as well as technical design choices, should prioritise those policies, choices and options that can help improve the position of those who are most disenfranchised within society.

[1226] Schemmel, 'Distributive and Relational Equality', 127; Pogge, 'Relational Conceptions of Justice', 2004, 147.

[1227] As Rawls stipulates: "The priority of liberty implies in practice that a basic liberty can be limited or denied solely for the-sake of one or more other basic liberties, and never, as I have said, for reasons of public good or of perfectionist values." Rawls, 'The Basic Liberties and Their Priority', 9.

EQUALITY OF CAPABILITY – Proponents of the capability approach believe that differences in people's advantage and privilege should be measured in terms of the substantial freedoms or real opportunities they have to choose between and realise (combined capabilities) alternative combinations of beings and doings (functionings). Substantial freedoms are created through a combination of personal abilities and suitable external conditions – shaped by the social, economic, political environment – that enable a person to choose and act.[1228]

Following a capability approach to algorithmic justice, the focus shifts towards those things people are able to do and be in life. Drawing on Nussbaum's list of central capabilities, I argued that the capabilities could be used as an evaluative and transformative benchmark to assess the (likely) impact algorithmic decision-making structures and acts, and the (in)equalities they produce, have had, or will have on people's substantial freedoms. They are the central substantial freedoms each person should have access to in order to lead a dignified life. The central capabilities can therefore indicate what should be considered salient within the algorithmic governance debate. More specifically, decision-makers should carefully consider whether their systems - once they have been placed within a wider social, economic, or socio-technical environment - are corrosive or fertile to people's overall opportunities and choice.

In addition, the capability approach demands the recognition of human diversity. Certain persons and groups face more challenges when they try to convert their resources into actual opportunities and functionings. In this respect, algorithmic structures and decisions should not become an institutional barrier that introduces and reinforces the entrenched injustice and inequality faced by certain members of society. Instead, these systems should be designed to drive social change, and where possible, accompanied by measures that enable each person to have their particular needs recognised and heard, and if necessary, accommodated.

LESSONS FOR EQUALITY AND NON-DISCRIMINATION LAW? – Through an analysis of the notions of oppression and domination, this research project defined how decision-subjects can experience status-based harms within the algorithmic environment. This articulation exercise could inform the evolutive interpretation of the legal equality concept. If socio-relational inequality is a concern shared by all - which European equality laws seem to recognise it is – the boundaries of the legal equality concept could be expanded so that every person can enjoy protection against status-based harms. Doing so might even be possible by modifying the existing language of non-discrimination law. For instance, the Courts' anti-stereotyping approach recognises how

[1228] Nussbaum, *Creating Capabilities: The Human Development Approach*, 20–21.

untruthful or prejudiced generalisations are often used to curtail the actions of (vulnerable) social groups. At the same time, judicial assessments based upon the vulnerability rationale acknowledge how certain groups due to their (historical) social position are more likely to become the subject of even greater socio-relational and/or distributive inequality. I have shown however how similar self-authorship-limiting constraints and vulnerabilities may be experienced by non-representational algorithmic groups as well. One should remain cautious however: while the harms experienced by both types of groups may share similarities, they are not identical, nor should they be conflated with one another. For instance, (members of) non-representational algorithmic aggregates may not share a history of disadvantage like certain social groups do. Decision-subjects differ in their lived experiences, which define the privilege, burdens and struggles each person carries individually. This human diversity must be recognised by the law and protection must be offered accordingly.

Lessons could also be learned from this project's examination of equality's distributive dimension. Whilst the identification of justice-relevant goods can help regulators identify or narrow down those domains of life where algorithmic decision-makers, due to their influence over them, hold a (greater) capacity to render decision-subjects vulnerable, regulatory initiatives could further widen the area in which this impact is to be monitored. Algorithmic decision-makers should be tasked to reflect upon the downstream impact their choices, actions, and decisions have on people's life prospects, even in those cases where they do not govern people's access to certain prized goods directly. Drawing on Creel and Hellman's notion of "realms of opportunity," the (initial) navigation of the digital environment could benefit from the identification of practices that are likely to affect people's (equal) access over (a broad range) of rights, liberties, opportunities, or capabilities. These decisions are more likely to have a greater impact on the socio-relational and socio-economic position of people and their life's future, and as a consequence, should be subject to review prior to their development and deployment.[1229] Following this view, decision-makers must reflect upon their position within an interconnected algorithmic environment – which in itself is part of a larger socio-technical, political, cultural and economic eco-system. If algorithmic inequality could be addressed regardless of where the (initial) interference took place or effect, the law would transcend its narrowly defined distributive boundaries.

COMPLEMENTARY? – Algorithmic acts of domination and oppression constrain decision-subjects in their access to certain justice-relevant goods. Likewise, when people are denied access to domains

[1229] See: Creel and Hellman, 'The Algorithmic Leviathan', 2021.

deemed critical for social and economic participation, they are rendered more vulnerable and hence more likely to become the subject of oppression and domination. In both cases, people are stripped from their desired degree of self-authorship. Though this points towards some level of complementarity between socio-relational and distributive ideals of equality, social egalitarians have rightfully observed how the application of a purely distributional logic onto non-material goods, such as power, opportunities and self-respect, can obscure how those aspects of our social life are the result of existing institutions, broader social structures and various other (interpersonal) relationships people may maintain.[1230] A society should not only focus on the things people should have or be able to do or be, but also regulate how institutions and people should treat each other. The viewpoints discussed are therefore not necessarily compatible. Albeit that the identified notions can be a start for regulators to deepen their understanding regarding the egalitarian harms algorithms are likely to produce, more efforts are needed to arrive at a complete theory of algorithmic justice that can confront both the socio-relational and distributive harms algorithms produce in a consistent and coherent manner, where strong socio-relational egalitarian guarantees are maintained through the support, and not only by virtue of a strong distributive safety net.

8.2.4 Part IV - Reconsidering the Concept of Equality for the Evaluation and Regulation of Algorithmically Guided Decision-Making.

Throughout this dissertation, I showed how a multidimensional understanding of equality can be operationalised as a guiding mechanism for the evaluation and regulation of algorithmic decision-making practices in light of the egalitarian harms they risk producing from cradle to grave. The regulatory potential of the concept of equality runs wide and deep. Informed by the identification, articulation, evaluation, and navigation exercises performed within this research project, my analysis of European equality and non-discrimination law and political philosophical theories of justice, I conclude this dissertation with a final set of recommendations to strengthen the position of equality as a social value within current algorithmic governance discourse. Suppose a law or legal instrument must be devised in order to review and assess the egalitarian harms of social and non-representational groups and their group members. What would such a regulatory review mechanism look like? This section formulates a cautious response thereto. I propose a review method in which the normative boundaries of equality can be further defined by positioning and balancing the egalitarian interests of decision-subjects, algorithmic groups, or society at large

[1230] Young, *Justice and the Politics of Difference*, 25–31.

against competing interests. In addition, I formulate recommendations on how this review could be structured, or more precisely, I pinpoint what elements substantially contribute to a meaningful evaluation and regulation of algorithmic decision-making systems.

Equality as Defined through a Balancing of Interests?

This research project has allowed for the articulation of various egalitarian objectives and harms. Still, the complexity of the algorithmic eco-system is not easily captured through a singular or fixed understanding of equality. Though we can identify and specify the societal ideals equality represents, whether those objectives can be achieved in practice, must be determined by a contextual analysis. We must learn to identify those algorithmic inequalities we are willing to tolerate. We must discover those inequalities that can be reasonably and objectively justified.

A CONTEXTUAL ANALYSIS - As this research project has shown, an investigation into the social impact of algorithms is affected by a multitude of factors. We must monitor how algorithms interact with our living environment over extended periods of time and investigate the specific impact algorithms will have in practice, taking into account the wider social, technical, economic, legal, and political culture in which these systems interact and operate. We must evaluate the legitimacy of the objectives pursued and assess whether the measure at hand does not exceed the limits of what is appropriate and necessary. We must carefully consider the choices made by decision-makers before, during and after the design, development and deployment of algorithmic systems, the power decision-makers have to impose those choices, and the way decision-makers make use of these powers. We must examine the domains in which algorithms will be deployed, and how one decision may interconnect with others, as part of larger algorithmic structures. We must consider how a decision may affect the (social) position of individual decision-subjects and groups of decision-subjects, and how those changes may affect society at large. We must understand how algorithms affect people's capacity to learn, develop, exercise, and express skills, capacities, and experiences in socially meaningful and recognised ways, or their capacity to exercise control over their actions and the conditions of their actions. We must consider the (institutional) resources, safeguards and structures, decision-subjects have been given to challenge unjust inequalities and pursue their fundamental interests. If all these factors shape our evaluation of algorithmic (in)equalities, we must have access to review mechanisms that enable all these contextual factors to be taken into consideration.

A classic legal approach where restrictions to equality are assessed via a balancing of (competing) interests can be taken as a point of departure.[1231] Under a balancing view, the (risk of) moral, socio-relational or distributional (or other) egalitarian harm produced, or induced by, a given algorithmic structure or differentiation act, and the disadvantage this may (or has) cause(d), is weighed against other competing interests, including the legitimate aims and interests pursued by the algorithmic decision-maker and the benefits those systems allegedly produce. An inquiry into the legitimacy, appropriateness and necessity of a measure may remain a valuable component of the review procedure. Still, even if it is the case that the aims pursued by a decision-maker are legitimate, the procedure chosen reasonably likely to achieve said aims, and no less restrictive means are available for realising them, it should still be assessed whether a decision-making act does not have an excessive effect on the (egalitarian) interests of those affected and society at large.

The wider environment and institutions in which these systems interact and operate serve as the background against which this balancing exercise will take place. Attention will therefore need to be paid to the unique properties algorithms have, such as their capacity to amplify or reinforce disadvantage, the danger of which can also be tied to the level of power a decision-maker holds. This also means that substantial and intelligible insight must be given into a system's functioning. In addition, (the risk of) harm may be offset by the procedural and institutional safeguards decision-subjects have been provided by the law or protective measures implemented by algorithmic decision-makers as far as decision-subjects also have the ability to make effective use of them. Once all relevant information is placed onto the balancing scale, a meaningful assessment can be made regarding the restrictions algorithms (risk to) impose on people's egalitarian interests, and ultimately, to gauge whether those restrictions can be justified.

PERFORMED THROUGHOUT A DECISION'S LIFE CYCLE - This balancing exercise must be performed throughout the entire decision-making chain. For decision-makers, it constitutes an ex-ante reflexive mechanism. Taking into account the egalitarian harms systems are likely to produce on an individual and societal level, decision-makers can assess whether the automation of a given decision-making process would be desirable in the first place. If so, they can start to reflect upon what design choices would be appropriate to mitigate potential egalitarian harm, and what

[1231] Though inspired by a traditional approach towards equality, the review process described is not identical. For example, certain elements of a classic (and formal) legal equality review, such as the comparator approach, would be discarded as a traditional comparative analysis would fail to capture the nuanced complexity of the algorithmic environment. See Section 8.2.2: "Part II - Equality and Non-Discrimination Law: Evaluating the Legal Conceptualisation of Equality and Algorithmic Injustice."

additional measures they could provide decision-subjects. After deployment, the balancing exercise can also act as an ex-post means to periodically re-examine the actual consequences of their system. At the same time, this approach can be used as a legal model for judicial authorities or regulators, or as a monitoring tool used by national authorities, such as external oversight boards. Here too, a balancing of interests should not be limited to a singular and isolated exercise. It should be performed periodically so as to map the likely impact a system may have over time, and if necessary, course-corrected in order to avoid further harm from happening. In other words, it should not only be used to evaluate harm once it has manifested.

THAT IS WELL-STRUCTURED, INCLUSIVE AND SUBJECT TO REVIEW – The equality review should be well-structured, coherent, and consistently applied as this would increase clarity and facilitate our ability to derive further guidance from them. Whomever performs this balancing exercise however, an additional condition is imposed onto it: it should always be possible to make the end-result and the components that were taken into account subject to an external review. This review process must be open. It should facilitate an inclusive, participatory, and deliberative, rather than an enclosed, public dialogue regarding the regulation of algorithmic decision-making, and the place equality holds therein. Sufficient attention must therefore be paid to the voices of those who often are or risk to be unheard or underrepresented. Their lived experiences should not only inform the performance of the review, but also its evaluation afterwards. By following such an approach, the goal is to eliminate as much as possible the risk that certain actors are given too much discretionary power in the conceptualisation of equality and could do so solely to their benefit or to the benefit of a majority. The goal would then be to reach a consensus regarding the various interests at play, and hence, incrementally built towards a more coordinated response on how to tackle algorithmic harms.

Motivating the Balancing Review

By performing this balancing exercise in a transparent and structured manner, we might come to an incremental understanding regarding the impact algorithms have on equality over time. This understanding not only concerns the core ideals we ascribe to equality as a structuring social value. It also relates to our ability to identify in practice those conditions under which those ideals are put under pressure, and the measures we identify as being most protective of them. By choosing to specify unjust inequalities through a balancing of interests however, the precise (normative) contours of equality are left somewhat undefined. Equality is not a rigid notion that represents a clearly defined set of measures and rules on how decision-makers should

accommodate decision-subjects egalitarian interests. Though this might give room for (regulatory) uncertainty, the choice for a balancing approach can be further motivated.

AN EVOLUTIVE, OPEN-ENDED APPROACH - First, while the notion of equality holds within it certain core goals or ideals, a balancing approach towards equality recognises that equality, rather than being a singular or fixed notion, is subject to (social) change. Our understanding of equality has evolved and will continue to evolve. Certain acts not (yet) considered an injustice today, will become so in the future (or vice versa). Technological changes may require us to re-assess the specific values we have assigned to equality, and the particular egalitarian harms we look to guard people against. When certain technological practices strike us as unjust, but we fail to articulate why exactly that is the case, we could reconceptualise our current understanding of equality and verify whether the response formulated by this renewed interpretation (continues to) correspond(s) to our intuitions. This dissertation can be placed within this tradition. Making use of my multidimensional model of equality, I have argued that the legal conceptualisation of equality is currently unfit to capture the egalitarian harms algorithms risk to impose. Still, this articulation process is never-ending, and, in an environment characterised by rapid technological advancements, such an exercise might need to be performed periodically. An open-ended approach recognises and accommodates the plural and dynamic nature of equality.

TACKLING COMPLEXITY - Second, a balancing exercise acknowledges the complexity of the algorithmic environment and the challenges these complexities pose for the monitoring and measurement of egalitarian harm. The transformative impact algorithmic technologies will have on the social fabric of society is often uncertain. Actual change may only become visible over time. As a consequence, it can be difficult to assess *a priori* how a given decision-making process will interact with its environment and whether it will produce an outcome that is undesirable. These challenges of unpredictability are furthermore compounded in cases where algorithmic decision-making acts are an integral component of larger socio-technical structures. This uncertainty can be mitigated by allowing a periodic balancing of interests. On an ex-ante level, a balancing review monitors whether the reasonably foreseeable egalitarian harms a system risks imposing can be outweighed by competing interests. On an ex-post level, a periodic balancing review can investigate whether that response would still be the same in light of the actual egalitarian harm a system has produced. In the case the response is different, the system should be abandoned, or additional safeguards to mitigate the harm should be proposed. Furthermore, it also allows egalitarian harms to be identified as they arise as a result of algorithmic structures, out of the actions of many decision-makers, rather than any given decision in isolation. This would require a birds-eye perspective that individual decision-makers might not have.

Complexity not only stems from uncertainty. The algorithmic environment comprises a multitude of actors who each have their own (competing) interests. Among those actors with competing interests are the decision-subjects themselves. The differentiation grounds determine which benefits or harms will be carried by a group, and therefore, the individuals that populate the group. The interests of the parties involved can diverge or converge on both levels interdependently and concurrently. This situation becomes even more complicated knowing that these groups are created continuously. Individuals thus partake in multiple groups, and potentially represent different interests, at the same time. Within a decision-making context, conflicting interests may exist (and change over time) a) between groups, taken as a whole b) between the members of one group vis-à-vis members of another group, c) between a member of a group and the group of which they are part and d) amongst members within the same group. The identified interests may furthermore diverge from the broader social interests a society may want to pursue when it looks to combat the societal changes algorithms may initiate.

While the dimensional model could be used to identify why a given disadvantage is inegalitarian, it might not succeed in formulating a concrete answer as to how these harms should be effectively addressed in highly complex situations. Yet, as part of an open-ended balancing review, we can start to formulate a beginning to such a response. For instance, by allowing the possibility to have both individual, as well as collective interest represented. Still, further efforts are needed to ensure the notion of equality can supersede the insular evaluation of decision-making practices and adequately consider the wider environment in which these practices, and their impact on people's interests, take place

ACCOMMODATING HARM - Third, within a balancing exercise, equality can capture the disadvantage of both (historically disadvantaged) social groups and non-representational algorithmic aggregates. On the one hand, for oppressed groups, the algorithmic environment acts as an additional institutional layer that blocks meaningful social participation. Under a balancing approach, equality could maintain its expressive and symbolic value, highlighting and prioritising the needs of those most disenfranchised. On the other hand, though members of algorithmic aggregates might not face structural or institutional constraints, they too might face difficulties in contesting disadvantage. Non-representational and seemingly random groups lack social cohesion. Because they are therefore more unlikely to organise themselves, they also lose out on their strength in numbers, having to resort to individualised claims for justice instead. In these cases, the precise social impact of algorithmic decision-making might also be more difficult to map. Under a balancing review, the various social positions and interactions decision-subjects may hold

as individuals and group members, may be better accounted for. The review procedure specifically requires that adequate consideration is given to a broad variety of contextual factors that shape a decision's (social) impact. By taking stock of how these factors might affect decision-subjects in different ways, differing measures could be proposed to accommodate their (diverging) needs.

RESEARCH LIMITATIONS - Finally, the proposed approach can accommodate the limitations of the research project. Though the abstract model of equality, and the various conceptualisations discussed, offer a representative view of the core goals equality aspires to achieve, as aforementioned, alternative conceptualisations exist. Societies can also be structured in non-egalitarian ways. As part of a balancing exercise, these alternative viewpoints and values are given the opportunity to compete with equality.
In sum, a balancing test allows for an evolutive and dynamic interpretation of equality. Furthermore, due to the additional condition that this exercise should be made subject to review, the interpretation and conceptualisation of equality itself can become subject to public scrutiny.

The Importance of Articulation

A balancing exercise will be fair when it is clearly articulated, well-structured and transparent. If the injustice of inequalities is defined via a balancing of interests, the goals and ideals we ascribe to equality, and the egalitarian harms we look to guard people against, should be clearly defined before they are positioned against competing interests. Accordingly, the articulation condition does not apply to equality alone, but to all interests involved and this lets people know what interests are at stake, and why they are worthy of protection.

Once an interest has been articulated, its position within the regulatory debate is strengthened and is more difficult to dismiss. Decision-makers will need to make their case in direct response to the interests involved. They would need to specify specifically why those interests can be overridden or, they would have to demonstrate that appropriate measures have been put in place so those interests could still be attained.

The articulation of interests may also help us understand alternative viewpoints. Furthermore, once a viewpoint is chosen over known and reasonable alternatives, this choice must be explained in relation to those alternatives. For instance, when decision-makers integrate a formal fairness metric into the design of their system, they would need to explain why this metric was chosen over others. Indeed, actors who hold power over the design and deployment of decision-making

systems should reflect upon the assumptions and values that underlie a decision-making process and justify their incorporation.

This articulation requirement also extends to those actors who will govern or evaluate the desirability of decision-making systems, be that the legislator, judiciary, or external evaluation boards. In this regard, current equality and non-discrimination case-law could further improve by explicating the normative underpinnings that ground the concept of equality. Especially when a supposed pluralist view of equality is taken, it is particularly important to understand which rationales are applied for the evaluation of differential treatment, when and why.

Articulation not only applies to the concepts at play, but also to the evaluation process itself. The review procedure must be clearly structured and consistently applied. This does not mean there is no room for change. As argued in the previous section, one benefit of a balancing review is its flexibility. We should be able to incorporate newfound information, and this in turn, may affect how algorithmic decision-making systems should be regulated. When this is the case however, these considerations must be explicated: what changes motivate an altered approach, and why? This clarity is necessary for persons to derive proper guidance from the balancing review.

Fundamental Interests as the Point of Departure

In our attempt to regulate algorithmic decision-making systems, we might be tempted to start our inquiry trying to identify those differences between people that make a difference, and subsequently assess whether the features relied upon by algorithms comprise an appropriate differentiation ground. While the evaluation of features and attributes would still be an element of our review, by focusing on them first, we would get the order wrong.

In light of the many differentiation grounds algorithms can produce, and the domains in which algorithms can find their purpose, the differences between individuals and groups of individuals that make a difference cannot be identified *in abstracto*. Differentiation grounds do not have an absolute nature that allows us to predict with certainty whether their use should be prohibited. For instance, a protected characteristic, such as a person's ethnicity, gender, or religion, could be used as a valid differentiating factor to realise full equality by means of positive action measures. The differentiation ground must therefore be assessed taking into account the context in which a decision is made and the risks that are involved for people's (egalitarian) fundamental interests. Taking these interests as a point of departure, guiding rules can be formulated on how the

algorithmic environment is best structured and governed so people's enjoyment of them can be maximised, rather than interfered with.

This shift to people's fundamental interests as our point of departure further provides an (initial) evaluative advantage. Even if not all information regarding a system's logic or functioning can be given, through an examination of the socio-relational and distributive consequences a system is likely to produce, an appraisal regarding their desirability could already be made *in part*. For example, when the exclusionary harms a system may introduce are too great, the deployment of the system should be reconsidered or made conditional upon the implementation of additional safety measures, such as the obligation for decision-makers to make their systems intelligible. Because we are concerned that a system might interfere with decision-subjects' fundamental egalitarian interests, we mandate transparency and explicability, including insight into the features that were identified as relevant by the decision-maker. In addition, and as aforementioned, certain decision-making practices such as incontestable generalisations are problematic irrespective of the generalisation's content or form simply because their use imposes onto decision-subjects a stand-alone socio-relational status-harm.

Of course, that does not mean that differentiation grounds do not matter in our evaluation of algorithmic decisions at all. Differentiation is inherent to decision-making, and in many cases, we cannot do without an evaluation of the differentiation grounds in question. If there is - based upon our initial contextual analysis, including an assessment of the egalitarian disadvantage a system is likely to generate for those involved - no immediate indication that a system is (socially) undesirable, we would still need to verify whether the differences identified and relied upon are effectively objective, reasonably justifiable and (relatively) connected to the purposes pursued by the decision-maker. Yet, here too, a clear articulation of egalitarian interests may inform decision-makers as to which differentiation grounds are more likely to be a justified point of differentiation. For example, the protected characteristics of non-discrimination laws could function as signifiers of morally reprehensible forms of differentiation and social disadvantage, whose use requires greater justification. In this regard, we are also encouraged to consider decision-making practices more holistically, rather than in truncated fashion. Even if the use of a given attribute or characteristic sounds valid or relevant at first sight, we might want to reconsider its use if it would appear in practice that it ultimately runs counter to the egalitarian objectives we wanted to achieve. For instance, this could be the case once we realise that a prima facie relevant ground was actually produced out of a biased data set. We could determine that the use of this attribute should now be prohibited, either out of concern that it would lead to unjust inequality in the future

or because we believe that the production of information via inegalitarian datasets in itself constitutes a wrong.

As every algorithmic differentiation ground captures an algorithmic group or aggregate, we would further need to know the data and attributes that were used to identify the individuals and groups that were effectively harmed by a system. This could also help us measure the actual harm involved. Put differently, whereas conceptualisations of equality can help us in identifying and articulating the (initial) risk of harm, the differentiation grounds, and the data used for their creation, can help us further specify who was subject to said disadvantage.

Equality: A Notion in Need of Support

The core of the balancing exercise pits the clearly articulated egalitarian interests of decision-subjects against those of algorithmic decision-makers. Both are positioned within the larger algorithmic environment, and the (compound) socio-relational and socio-economic inequalities they might impose.

If the egalitarian interests of decision-subjects are to be taken seriously, those interests are in need of support: decision-subjects must be empowered and protected in order to ensure they can enjoy their fundamental interests with equal intensity. Throughout this dissertation, in both the legal and political philosophical analysis, the importance of measures associated with institutional and procedural fairness, as well as mechanisms enabling inclusion, participation and control, came to the forefront. Such safeguards serve a threefold function. First, they make it possible to assess the (likely) egalitarian impact of algorithmic decision-making systems. Second, they resource and protect decision-subjects against suffering egalitarian harm. Third, they are a necessary condition if technologies are to become a driver for change that can actively break down institutional barriers towards equality. These safeguards should therefore feature prominently in the toolbox of decision-makers and regulators. An assessment of their presence should be a key point of attention in the evaluation of algorithmic decision-making systems and the regulation thereof

In order to ensure a meaningful evaluation of decision-making systems, their functioning should be made transparent, traceable, identifiable, and explicable. Transparency not only applies to the data used and the logic involved in the derivation and application of data-driven knowledge. It also pertains to the choices that were made leading up to the design and development of automated systems. Why was automation deemed desirable in the first place? Similarly, in order

to map the extended impact algorithms might have over time and across domains, it must also be explained to what extent a particular decision builds upon or feeds into other decisions. Likewise, insight must be given regarding the position an individual system may within larger technological and societal structures. This information not only facilitates the identification and measurement of egalitarian harms, or the risk thereof, it also makes it possible to verify whether the aims and motivations pursued by the decision-maker can be achieved in practice in a manner that is reliable and sound. Other measures of accountability, such as reflexive self-monitoring and documentation obligations, as well as data management strategies, could be key in the protection of equality.[1232] The latter mechanisms necessitate decision-makers to reflect upon the assumptions that underlie their decision-making processes, and the estimated impact their systems might have on others. Yet, they are also a prerequisite to make their interests explicit, visible, and hence reviewable. Though mechanisms of accountability and transparency in themselves are unlikely to be sufficient for breaking down the informational knowledge gap that exists between decision-subject and decision-maker - nor do they necessarily render the complexity of the algorithmic environment comprehensible - their presence can be seen as a necessary step in giving decision-subjects a better chance to navigate the digital environment in ways that are foreseeable and deliberate.

Mechanisms of information provision, access, control, and contestation - which could be provided in the form of individual (or collective) data rights – could further enhance the position of decision-subjects. They can give decision-subjects the ability to foresee, assess and challenge the inegalitarian nature of a decision-making process, as well as give them more certainty, autonomy, and voice. The promotion of diversity and inclusion, among others through participatory design strategies, could further ensure that the development, deployment, and monitoring of decision-making systems is confronted by a plurality of views, and informed by the lived experiences of others. Institutional and procedural safeguards give decision-subjects the ability to equalise power asymmetries within the digital environment and help them regain control over those structures that govern and shape their choices, options, and conditions of their actions in ways that are heard and seen in socially meaningful and recognised ways. They can therefore contribute to the democratisation of the modern means of interpretation and communication, which algorithms have become, and help realise the transformative function technologies could have in breaking down barriers. Even more so, if equality is defined as the need to secure each person

[1232] Timnit Gebru et al., 'Datasheets for Datasets', *ArXiv:1803.09010 [Cs]*, 19 March 2020, http://arxiv.org/abs/1803.09010; Miceli et al., 'Documenting Computer Vision Datasets'.

access to self-authorship, self-determination and self-development, the presence (and enforcement) of the aforementioned measures becomes a precondition to achieve equality in an algorithmically mediated environment.

As the algorithmic environment is comprised by great complexity and the egalitarian harms systems produce are often imposed onto group-level structures, one should nonetheless be cautious in placing too much responsibility onto the individual decision-subject. Individual mechanisms for review, control, and participation that decision-subjects should have access to in order to realise their fundamental interest to self-determination and self-development, must be combined with additional (legal) mechanisms so that other actors, such as regulators or civil society, can help them in their cause. Furthermore, collective interest groups could also play a role in addressing the social or collective harms decision-making processes might generate on a structural or aggregate level, and if need be, weigh these social or collective interests against the diverging interests individual decision-subjects and decision-makers may have.
From a research and regulatory perspective, the complementary nature between (substantive) equality and institutional, procedural or participatory safeguards - and the presence of such measures in the current regulatory toolbox decision-subjects and interests groups have to their availability – should be further explored in order to identify what additional regulatory measures could help realise the goals of equality as a positive and transformative ideal.[1233]

[1233] Some notable examples of scholarship that have already tried to explore this relationship can be found in: Viljoen, 'A Relational Theory of Data Governance'; Gerards et al., *Algorithmic Discrimination in Europe*; Wachter, 'The Theory of Artificial Immutability'. Fisher and Streinz, 'Confronting Data Inequality'. Miceli et al., 'Documenting Computer Vision Datasets'. Abeba Birhane et al., 'Power to the People? Opportunities and Challenges for Participatory AI', 15 September 2022, https://doi.org/10.1145/3551624.3555290.

Bibliography

Laws

- Council of Europe, European Convention for the Protection of Human Rights and Fundamental Freedoms, as amended by Protocols Nos. 11 and 14, 4 November 1950, ETS 5, available via: <https://www.refworld.org/docid/3ae6b3b04.html>.

- Council of Europe, Protocol 12 to the European Convention on Human Rights and Fundamental Freedoms on the Prohibition of Discrimination, 4 November 2000, ETS 177, available via: <https://www.refworld.org/docid/3ddd0cb44.html>.

- Consolidated Version of the Treaty on the Functioning of the European Union (2012) OJ C-326/47 (TFEU); available via: <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=celex%3A12012E%2FTXT>.

- Consolidated Version of the Treaty on European Union (2012) OJ C-326/13 (TEU); available via: <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A12012M%2FTXT>.

- Charter of Fundamental Rights of the European Union (2000) OJ C-361-1 (CFR); available via: <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=celex%3A12012P%2FTXT>.

- Explanations relating to the Charter of Fundamental Rights (2007) OJ C 303/17; available via: <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:32007X1214(01)>.

- Proposal for a Regulation of the European Parliament and of the Council Laying Down Harmonised Rules on Artificial Intelligence (Artificial Intelligence Act) and Amending Certain Union Legislative Acts (COM/2021/206 Final); available via: <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A52021PC0206>.

- Proposal for a Regulation of the European parliament and of the Council on a Single Market for Digital Services (Digital Services Act) and amending Directive 2000/31/EC (Com/2020/825 Final); available via: https://eur-lex.europa.eu/legal-content/en/TXT/?qid=1608117147218&uri=COM%3A2020%3A825%3AFIN

- European Parliament resolution of 20 October 2020 with recommendations to the Commission on a framework of ethical aspects of artificial intelligence, robotics and related technologies (2020/2012(INL); 2021/C 404/04); available via: <https://eur-lex.europa.eu/legal-content/EN/TXT/HTML/?uri=CELEX:52020IP0275>.

- Directive (EU) 2016/681 of the European Parliament and of the Council of 27 April 2016 on the use of passenger name record (PNR) data for the prevention, detection, investigation and prosecution of terrorist offences and serious crime; *OJ L 119, 4.5.2016*; available via: <https://eur-lex.europa.eu/eli/dir/2016/681/oj>.

- Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and

repealing Directive 95/46/EC (General Data Protection Regulation) (Text with EEA relevance), available via: <https://eur-lex.europa.eu/eli/reg/2016/679/oj>.

- Directive 2002/58/EC of the European Parliament and of the Council of 12 July 2002 concerning the processing of personal data and the protection of privacy in the electronic communications sector (2002) OJ L201; available via: <https://eur-lex.europa.eu/legal-content/EN/ALL/?uri=CELEX%3A32002L0058>.

- Proposal for a Council Directive on implementing the principle of equal treatment between persons irrespective of religion or belief, disability, age or sexual orientation {SEC(2008) 2180} {SEC(2008) 2181}; available via: <https://eur-lex.europa.eu/legal-content/en/TXT/?uri=CELEX%3A52008PC0426>.

- Directive 2006/54/EC of the European Parliament and of the Council of 5 July 2006 on the implementation of the principle of equal opportunities and equal treatment of men and women in matters of employment and occupation (recast); Available via: <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A32006L0054>.

- Council Directive 2004/113/EC of 13 December 2004 implementing the principle of equal treatment between men and women in the access to and supply of goods and services; Available via: <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=celex:32004L0113>.

- Council Directive 2000/78/EC of 27 November 2000 establishing a general framework for equal treatment in employment and occupation; Available via: <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=celex%3A32000L0078>.

- Council Directive 2000/43/EC of 29 June 2000 implementing the principle of equal treatment between persons irrespective of racial or ethnic origin (Racial Equality Directive); Available via: <https://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=CELEX:32000L0043:en:HTML>.

Case-Law

European Court of Human Rights

- ECtHR Grand Chamber (15 June 2021) Case of Kurt v. Austria (Application no. 62903/15)
- ECtHR (4 May 2021) Case of Jurčić v. Croatia, Application no. 54711/15.
- ECtHR First Section (24 October 2019, final 24 February 2020), Case of J.D. and A. v. The United Kingdom, Application nos. 32949/17 and 34614/17.
- ECtHR Fourth Section (19 February 2019), Case of Aleksić v. Slovenia, Application no. 57123/10.
- ECtHR Fourth Section (16 April 2019) Case of Lingurar v. Romania, Application no, 48474/14.
- ECtHR First Section (13 September 2018) Case of Big Brother Watch and Others v The United Kingdom App nos 58170/13, 62322/14 and 24960/15.
- ECtHR Grand Chamber (19 December 2018) Case of Molla Sali v Greece, Application no. 20452/14.
- ECtHR Grand Chamber (5 September 2017), Case of Fabian v. Hungary, Application no. 78117/13.

- ECtHR (25 July 2017, Final 25 October 2017) Case of Carvalho Pinto de Sousa Morais v. Portugal, Application no. 17484/15.
- ECtHR (11 July 2017), Case of Dakir v. Belgium, Application no. 4619/12.
- ECtHR Fifth Section (23 March 2017, Final 23 June 2017), Case of Wolter and Sarfert v. Germany, Application nos. 59752/13 and 66277/13.
- ECtHR (21 July 2016, 30 January 2017 Final) Case of Mamatas and Others v. Greece, Application nos. 63066/14, 64297/14 and 66106/14.
- in ECtHR Grand Chamber (24 May 2016) Case of Biao v. Denmark, Application no. 38590/10
- ECtHR (15 September 2016, Final 30 January 2017) Case of British Gurtkha Welfare Society and Others v. the United Kingdom.
- ECtHR (16 November 2014), Case of Ünal Tekeli v. Turkey, Application no. 29865/96.
- ECtHR (22 November 2012; Final 29 April 2013) Case of Cadek and Others v. The Czech Republic, Applications nos. 31933/08, 60084/08, 6185/09, 46696/09, 52792/09, 53518/09, 10185/10, 42151/10, 3167/11 and 20939/11.
- ECtHR (24 July 2012, Final 24 October 2012) Case of B.S. v. Spain, Application no. 47159/08
- ECtHR Grand Chamber (22 March 2012) Case of Konstantin Markin v. Russia, Application no. 30078/06.
- ECtHR Fourth Section (27 September 2011, Final 27 December 2011), Case of Bah v. The United Kingdom, Application no. 56328/07.
- ECtHR First Section (10 March 2011, Final 15 September 2011), Case of Kiyutin v. Russia, Application no. 2700/10.
- ECtHR (4 May 2010) Josef Peterka v. Czech Republic, Application no. 21990/08.
- ECHR (27 April 2010) Springett and Others v United Kingdom, Application No. 34726/04, 14287/05 and 34702/05 (Admissibility).
- ECtHR (13 July 2010, 22 November 2010, Final) Case of Clift v.The United Kingdom, Application no. 7205/07.
- ECtHR (16 March 2010) Case of Oršuš and Others v. Croatia, Application no. 15766/03.
- ECtHR Grand Chamber (16 March 2010) Case of Carson and Others v. The United Kingdom, Application no. 42184/05.
- ECtHR Fourth Section (2 March 2010) Case of Kozak v. Poland, Application no. 13102/02.
- ECtHR Grand Chamber (12 February 2008), Case of Kafkaris v. Cyprus, Application no. 21906/04.
- ECtHR Grand Chamber (13 November 2007) Case of D.H. and Others v. The Czech Republic, Application no. 57325/00.
- ECtHR Second Section (13 December 2005, final 13 March 2006), Case of Timishev v. Russia, Application nos. 55762/00 and 55974/00.
- ECtHR Grand Chamber (6 July 2005) Case of Nachova and Others v. Bulgaria, Application nos. 43577/98 and 43579/98.
- ECtHR Grand Chamber (11 July 2002), Case of Christine Goodwin v. The United Kingdom (Application no. 28957/95).
- ECtHR Grand Chamber (28 May 2002), Case of Stafford v. the United Kingdom, Application no. 46295/99.
- ECtHR Grand Chamber (10 May 2001) Case of Cyprus v. Turkey, Application no. 25781/94.
- ECtHR (20 June 2000) Case of Magee v the United Kingdom App no. 28135/95.
- ECtHR Chamber (22 February 1994), Case of Burghartz v. Switzerland, Application no. 16213/90.
- ECtHR (28 November 1984) Case of Rasmussen v. Denmark, Application no. 8777/79.
- ECtHR Chamber (9 October 1979), Case of Airey v. Ireland, Application no. 6289/73.
- ECtHR (7 December 1976), Case of Handyside v. The United Kingdom, Application no. 5493/72.
- ECtHR (7 December 1976) Case of Kjeldsen, Busk Madsen and Pedersen v. Denmark, Application no. 5095/71; 5920/72; 5926/72.

- ECtHR (8 June 1976) Case of Engel and Others v. The Netherlands, Application no. 5100/71; 5101/71; 5102/71; 5354/72; 5370/72.
- European Commission on Human Rights (14 December 1973) East African Asians v. the United Kingdom, Applications nos. 4403/70-4419/70, 4422/70, 4423/70, 4434/70, 443/70, 4476/70-4478/70, 4486/70, 4501/70 and 4526/70-4530/70.
- ECtHR (23 July 1968) Case "Relating to Certain Aspects of the Laws on the Use of Languages in Education in Belgium" v. Belgium (Merits)," App no 1474/62, 1677/62, 1769/63, 1994/63, 2126/64.

European Court of Justice

- CJEU (21 June 2022), Case C-817/19 Ligue des Droits Humains v. Conseils des Ministres; ECLI:EU:C:2022:491
- CJEU (5 April 2022) Case C-140/20 G.D. v. Commissioner of An Garda Siochana; ECLI:EU:C:2022:258
- CJEU (15 July 2021) Joined Cases C-804/18 and C-341/19 IX v. WABE EV and MH Müller Handels GmbH v. MJ, ECLI:EU:C:2021:594
- CJEU (26 January 2021) Case C-16/19 VL v. Szpital Kliniczny im. dra J. Babińskiego Samodzielny Publiczny Zakład Opieki Zdrowotnej w Krakowie, ECLI:EU:C:2021:64
- CJEU (21 January 2021) Case-843/19 of INSS v. BT, ECLI:EU:C:2021:55.
- CJEU (29 October 2020), Case C-243/19 A v Veselības ministrija; ECLI:EU:C:2020:872
- CJEU (6 October 2020) La Quadrature du Net and Others v Premier Ministre and Others; ECLI:EU:C:2020:791
- CJEU (17 April 2018) Case C-414/16 Egenberger v. Evangelisches Werk Fur Diakonie und Entwicklung eV; ECLI:EU:C:2018:257
- CJEU (18 January 2018) Case C-270/16 of Carlos Enrique Ruiz Conejero v. Ferroser Servicios Auxiliaires SA, Ministerio Fiscal, ECLI:EU:C:2018:17
- CJEU (26 July 2017) Opinion 1/15 of the Court on the transfer of Passenger Name Record Data from the European Union to Canada, ECLI:EU:C:2017:592
- CJEU (5 July 2017) Case C-190/16 Werner Fries v. Lufthansa CityLine GmbH, ECLI:EU:C:2017:513
- CJEU (14 March 2017) Case C-157/15 Samira Achbita and Centrum voor Gelijkheid van Kansen en voor Racismebestrijding v. G4S Secure Solutions, ECLI:EU:C:2017:203.
- CJEU (9 March 2017) C-406/15 of Milkova; ECLI:EU:C:2017:198
- CJEU (21 December 2016) Joined Cases C-203/15 and C-698/15, Tele2 Sverige AB, ECLI:EU:C:2016:970
- CJEU (24 November 2016) Case C-443/15 of Parris v. Trinity College Dubline (and others), ECLI:EU:C:2016:897.
- CJEU (13 July 2016) Case C-188/15 Asma Bougnaiou and Association de Defense des Droits de l'Homme v. Micropole SA, ECLI:EU:C:2017:204.
- CJEU (16 July 2015) Case C-83/14 Chez, ECLI:EU:C:2015:480.
- CJEU (29 April 2015) Case C-528/13 Léger, ECLI:EU:C:2015:288
- CJEU (18 December 2014) Case C-354/13 of Fag og Arbejde (FOA) v. Kommunernes Landsforening (KL) *(Kaltoft),* ECLI:EU:C:2014:2463.
- CJEU (22 May 2014), Case C-356/12 Wolfgang Glatzel v. Freistaat Bayern, ECLI:EU:C:2014:350.
- CJEU (1 March 2011) Case C-236/09 Test-Achats; ECLI:EU:C:2011:100.
- CJEU (16 September 2010) Case C-149/10 Zoi Chatzi v Ypourgos Oikonomikon; ECLI:EU:C:2010:534
- CJEU (21 July 2011), Case C-21/10 of Károly Nagy v. Mezőgazdasági és Vidékfejlesztési Hivatal, ECLI:EU:C:2011:505.

- CJEU (16 December 2008) Case C-127/07 Société Arcelor Atlantique et Lorraine and Others v Premier ministre, Ministre de l'Écologie et du Développement durable and Ministre de l'Économie, des Finances et de l'Industrie ; ECLI:EU:C:2008:728
- CJEU (17 July 2008) Case C-303/06 Coleman, ECLI:EU:C:2008:415.
- CJEU (10 July 2008) Case C-54/07 Centrum voor Gelijkheid van kansen en voor racismebestrijding v. Firma Feryn NV, ECLI:EU:C:2008:397
- Opinion of AG Sharpston (22 May 2008) Case-427/06 Birgit Bartsch v Bosch und Siemens Hausgeräte (BSH) Altersfürsorge GmbH; ECLI:EU:C:2008:297.
- CJEU (22 November 2006) Case C-144/04 Werner Mangold v. Rüdiger Helm, ECLI:EU:C:2005:709
- CJEU (19 March 2002) Case C-476/99 H. Lommers v. Minister van Landbouw, Natuurbeheer en Visserij, ECLI:EU:C:2002:183.
- CJEU (13 April 2000) Case C-292/97, Kjell Karlsson and Others; ECLI:EU:C:2000:202
- CJEU (9 February 1999) Case C-167/97 Seymour-Smith, ECLI:EU:C:1999:60
- CJEU (11 November 1997) Case C-409/95 of Helmut Marschall v. Land Nordrhein-Westfalen, ECLI:EU:C:1997:533.
- CJEU (30 April 1996), Case C-13/94 P v S and Cornwall County Council; ECLI:EU:C:1996:170
- CJEU (17 October 1995) Case C-450/93 of Eckhard Kalanke v. Freie Hansestadt Bremen, ECLI:EU:C:1995:322.
- CJEU (8 November 1990) Case C-177/88 Dekker, ECLI:EU:C:1990:383.
- CJEU (13 May 1986) Case C-170/84 of Bilka - Kaufhaus GmbH v Karin Weber von Hartz**,** ECLI:EU:C:1986:204
- Case C 184/83 (12 July 1984) Ulrich Hofmann v Barmer Ersatzkasse, ECLI:EU:C:1984:273

Journals, Books, Reports and Miscellaneous

- 'A Kantian Conception of Equality by John Rawls '43 *50'. Accessed 17 August 2020. https://princetonindependent.com/issue01.03/item10c.html.
- Ad Hoc Committee on Artificial Intelligence (CAHAI). 'Feasibility Study'. Council of Europe, 17 December 2020.
- Ad Hoc Committee on Artificial Intelligence (CAHAI). 'Possible elements of a legal framework on artificial intelligence, based on the Council of Europe's standards on human rights, democracy and the rule of law', 3 December 2021; available via: https://rm.coe.int/cahai-2021-09rev-elements/1680a6d90d.
- 'AI's New Workforce: The Data-Labelling Industry Spreads Globally | Financial Times'. Accessed 12 May 2022. https://www.ft.com/content/56dde36c-aa40-11e9-984c-fac8325aaa04.
- Alexander, John M. *Capabilities and Social Justice: The Political Philosophy of Amartya Sen and Martha Nussbaum*. Ashgate, 2008.
- Alexander, John M. 'Capability Egalitarianism and Moral Selfhood'. *Ethical Perspectives*, no. 1 (2003): 3–21. https://doi.org/10.2143/EP.10.1.503868.
- Alexander, John M. 'Non-Reductionist Naturalism: Nussbaum between Aristotle and Hume'. *Res Publica* 11, no. 2 (June 2005): 157–83. http://dx.doi.org.kuleuven.ezproxy.kuleuven.be/10.1007/s11158-005-1483-6.
- ———. 'Social Justice and Nussbaum's Concept of the Person*'. In *Capabilities, Gender, Equality: Towards Fundamental Entitlements*, edited by Flavio Comim and Martha C. Nussbaum, 414–36. Cambridge: Cambridge University Press, 2014. https://doi.org/10.1017/CBO9781139059138.021.
- Alexander, Larry A. 'Fair Equality of Opportunity: John Rawls' Forgotten Principle'. *Philosophy Research Archives* 11 (1985): 197–208. https://doi.org/pra19851111.

- Allhutter, Doris, Florian Cech, Fabian Fischer, Gabriel Grill, and Astrid Mager. 'Algorithmic Profiling of Job Seekers in Austria: How Austerity Politics Are Made Effective'. *Frontiers in Big Data* 3 (2020): 5. https://doi.org/10.3389/fdata.2020.00005.
- Allo, Patrick. 'The Epistemology of Non-Distributive Profiles'. *Philosophy & Technology* 33, no. 3 (1 September 2020): 379–409. https://doi.org/10.1007/s13347-019-00360-z.
- Alvarado, Oscar, Elias Storms, David Geerts, and Katrien Verbert. 'Foregrounding Algorithms: Preparing Users for Co-Design with Sensitizing Activities', 2020, 7.
- Ananny, Mike. 'Toward an Ethics of Algorithms: Convening, Observation, Probability, and Timeliness'. *Science, Technology, & Human Values* 41, no. 1 (January 2016): 93–117. https://doi.org/10.1177/0162243915606523.
- Anderson, Elizabeth S. 'What Is the Point of Equality?' *Ethics* 109, no. 2 (1999): 287–337. https://doi.org/10.1086/233897.
- Angwin, Julia, Jeff Larson, Surya Mattu, Lauren Kirchner, and ProPublica. 'Machine Bias'. ProPublica, 23 May 2016. https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing?token=lZ_nPrh6oVJEnMzcTH1Jr59Ibe3K8XZC.
- Arnardottir, O. M. 'The Differences That Make a Difference: Recent Developments on the Discrimination Grounds and the Margin of Appreciation under Article 14 of the European Convention on Human Rights'. *Human Rights Law Review* 14, no. 4 (1 December 2014): 647–70. https://doi.org/10.1093/hrlr/ngu025.
- Arnardóttir, Oddný Mjöll. 'Discrimination as a Magnifying Lens: Scope and Ambit under Article 14 and Protocol No. 12'. In *Shaping Rights in the ECHR: The Role of the European Court of Human Rights in Determining the Scope of Human Rights*, edited by Eva Brems and Janneke Gerards, 330–49. Cambridge: Cambridge University Press, 2014. https://doi.org/10.1017/CBO9781107337923.020.
- ———. *Equality and Non-Discrimination under the European Convention on Human Rights*. Brill Nijhoff, 2021. http://brill.com/view/title/8988.
- ———. 'Multidimensional Equality from within: Themes from the European Convention on Human Rights.' In *European Union Non-Discrimination Law: Comparative Perspectives on Multidimensional Equality Law.*, edited by Dagmar Schiek and Victoria Chege, 53–72. London and New York: Routledge-Cavendish, 2009.
- ———. 'Rethinking the Two Margins of Appreciation'. *European Constitutional Law Review* 12, no. 1 (May 2016): 27–53. https://doi.org/10.1017/S1574019616000018.
- ———. 'The Differences That Make a Difference: Recent Developments on the Discrimination Grounds and the Margin of Appreciation under Article 14 of the European Convention on Human Rights'. *Human Rights Law Review* 14, no. 4 (1 December 2014): 647–70. https://doi.org/10.1093/hrlr/ngu025.
- ———. 'Vulnerability under Article 14 of the European Convention on Human Rights'. *Oslo Law Review* 4, no. 03 (15 December 2017): 150–71. https://doi.org/10.18261/issn.2387-3299-2017-03-03.
- Arneson, Richard. 'Equality of Opportunity'. In *The Stanford Encyclopedia of Philosophy*, edited by Edward N. Zalta, Summer 2015. Metaphysics Research Lab, Stanford University, 2015. https://plato.stanford.edu/archives/sum2015/entries/equal-opportunity/.
- Arneson, Richard J. 'Against Rawlsian Equality of Opportunity'. *Philosophical Studies: An International Journal for Philosophy in the Analytic Tradition* 93, no. 1 (1999): 77–112.
- ———. 'Equality and Equal Opportunity for Welfare'. *Philosophical Studies: An International Journal for Philosophy in the Analytic Tradition* 56, no. 1 (1989): 77–93.
- Article 29 Data Protection Working Party. 'Guidelines on Automated Individual Decision-Makinng and Profiling for the Purposes of Regulation 2016/79'. Article 29 Data Protection Working Party, 2018.
- Asher-Shapiro, Avi. 'Explainer - Exam Grading Algorithms amid Coronavirus: What's the Row About?' *Reuters*, 12 August 2020, sec. Healthcare. https://www.reuters.com/article/global-tech-education-idUSL8N2FD0EI.

- Atrey, Shreya. 'Beyond Universality : An Intersectional Justification of Human Rights'. In *Intersectionality and Human Rights Law*, edited by Shreya Atrey and Peter Dunne, 1st ed., 17–38. Oxford: Hart Publishing, 2020. http://www.bloomsburycollections.com/book/intersectionality-and-human-rights-law/ch1-beyond-universality/.
- ———. *Intersectional Discrimination*. Oxford: Oxford University Press, 2019. https://doi.org/10.1093/oso/9780198848950.001.0001.
- ———. *Intersectional Discrimination*. *Intersectional Discrimination*. Oxford University Press. Accessed 8 July 2020. http://www.oxfordscholarship.com/view/10.1093/oso/9780198848950.001.0001/oso-9780198848950.
- ———. *Intersectional Discrimination, Shreya Atrey.* Oxford, England: Oxford University Press, 2019.
- Atrey, Shreya, and Peter Dunne, eds. *Intersectionality and Human Rights Law*. Hart Publishing, 2020. https://doi.org/10.5040/9781509935321.
- 'Automating Society 2020'. AlgorithmWatch and BertelsmannSiftung, January 2021. https://automatingsociety.algorithmwatch.org/wp-content/uploads/2021/01/Automating_Society_Report_2020_-_Deutsche_Ausgabe.pdf.
- Baker, Aaron. 'Proportional, Not Strict, Scrutiny: Against a U.S. "Suspect Classifications" Model under Article 14 ECHR in the U.K.' *American Journal of Comparative Law* 56, no. 4 (1 September 2008): 847–94. https://doi.org/10.5131/ajcl.2007.0030.
- Baker, John. 'Conceptions and Dimensions of Social Equality'. In *Social Equality: On What It Means to Be Equals*, edited by Carina Fourie, Fabian Schuppert, and Ivo Walliman-Helmer, 65–86. New York: Oxford University Press, 2015.
- Balayn, Agathe, and Seda Gurses. 'Beyond Debiasing: Regulating AI and Its Inequalities.' EDRi, 2021.
- Balkin, J. M. 'The Constitution of Status'. *The Yale Law Journal* 106, no. 8 (June 1997): 2313. https://doi.org/10.2307/797222.
- Bamforth, Nicholas. 'Approaching the Indirect–Direct Discrimination Distinction: Concepts, Justifications and Policies'. In *Foundations of Indirect Discrimination Law*, edited by Hugh Collins and Tarunabh Khaitan, 1st ed., 57–82. Oxford: Hart Publishing, 2018. http://www.bloomsburycollections.com/book/foundations-of-indirect-discrimination-law/ch3-approaching-the-indirect-direct-discrimination-distinction-concepts-justifications-and-policies/.
- ———. 'Prohibited Grounds of Discrimination under EU Law and the European Convention on Human Rights: Problems of Contrast and Overlap'. *Cambridge Yearbook of European Legal Studies* 9 (ed 2007): 1–42. https://doi.org/10.5235/152888712802746920.
- ———. 'Sexual Orientation Discrimination after Grant v South-West Trains'. *Modern Law Review* 63, no. 5 (September 2000): 694–720. https://doi.org/10.1111/1468-2230.00287.
- Barman-Adhikari, Anamika, Eric Rice, Kimberly Bender, Rebecca Lengnick-Hall, Amanda Yoshioka-Maxwell, and Harmony Rhoades. 'Social Networking Technology Use and Engagement in HIV Related Risk and Protective Behaviors among Homeless Youth'. *Journal of Health Communication* 21, no. 7 (July 2016): 809–17. https://doi.org/10.1080/10810730.2016.1177139.
- Barocas, Solon, Moritz Hardt, and Arvind Narayanan. 'Fairness and Machine Learning', n.d., 253.
- ———. 'Fairness in Machine Learning', n.d., 127.
- Barry, Brian. *Theories of Justice: A Treatise on Social Justice, Volume I*. Berkeley and Los Angeles, California: University of California Press, 1989.
- Barry, Nicholas. 'Defending Luck Egalitarianism'. *Journal of Applied Philosophy* 23, no. 1 (2006): 89–107. https://doi.org/10.1111/j.1468-5930.2006.00322.x.

- BBC News. 'A-Levels and GCSEs: How Did the Exam Algorithm Work?' *BBC News*, 20 August 2020, sec. Explainers. https://www.bbc.com/news/explainers-53807730.
- Beeghly, Erin. 'Discrimination & Disrespect'. In *Routledge Handbook to the Ethics of Discrimination*, edited by Kasper Lippert-Rasmussen, 83–96. New York, NY: Routledge, 2017.
- Beer, David. 'The Social Power of Algorithms'. *Information, Communication & Society* 20, no. 1 (2017): 1–13. https://doi.org/10.1080/1369118X.2016.1216147.
- Bekkum, Marvin van, and Frederik Zuiderveen Borgesius. 'Digital Welfare Fraud Detection and the Dutch SyRI Judgment'. *European Journal of Social Security*, 2 August 2021, 13882627211031256. https://doi.org/10.1177/13882627211031257.
- Belavusau, Uladzislau, and Kristin Henrard. 'A Bird's Eye View on EU Anti-Discrimination Law: The Impact of the 2000 Equality Directives'. *German Law Journal* 20, no. 5 (July 2019): 614–36. https://doi.org/10.1017/glj.2019.53.
- ———. 'The Impact of the 2000 Equality Directives on EU Anti-Discrimination Law : Achievements and Pitfalls'. In *EU Anti-Discrimination Law Beyond Gender*, edited by Uladzislau Belavusau and Kristin Henrard, 1st ed., 1–38. Oxford,: Hart Publishing, 2019. http://www.bloomsburycollections.com/book/eu-anti-discrimination-law-beyond-gender/ch1-the-impact-of-the-2000-equality-directives-on-eu-anti-discrimination-law/.
- Bell, Mark. 'Article 20: Equality Before the Law'. In *The EU Charter of Fundamental Rights: A Commentary*, edited by Steve Peers, Tamara Hervey, Jeff Kenner, and Angela Ward, 563–78. Hart Publishing, 2014.
- ———. 'Article 20: Equality Before the Law'. In *The EU Charter of Fundamental Rights: A Commentary*, edited by Steve Peers, Tamara Hervey, Jeff Kenner, and Angela Ward, 563–78. Hart Publishing, 2021.
- ———. 'The Transformation of EU Anti-Discrimination Law'. In *Anti-Discrimination Law and the European Union*. Oxford: Oxford University Press, 2002. https://doi.org/10.1093/acprof:oso/9780199244508.003.0008.
- Bell, Mark, and Lisa Waddington. 'Similar Yet Different: The Work-Life Balance Directive and the Expanding Frontiers of EU Non-Discrimination Law.' *Common Market Law Review* 58, no. 5 (2021): 32.
- Benedi Lahuerta, Sara. 'Ethnic Discrimination, Discrimination by Association and the Roma Community: CHEZ'. *Common Market Law Review* 53, no. 3 (2016): 797–818.
- Ben-Israel, Isaac, Jorge Cerdio, Arisa Ema, Leehe Friedman, Marcello Ienca, Alessandro Mantelero, Eviatar Matania, Catelijne Muller, Hideaki Shiroyama, and Effy Vayena. "Towards Regulation of AI Systems: Global Perspectives on the Development of a Legal Framework on Artificial Intelligence Systems Based on the Council of Europe's Standards on Human Rights, Democracy and the Rule of Law." Council of Europe Study. Council of Europe, December 2020. https://rm.coe.int/prems-107320-gbr-2018-compli-cahai-couv-texte-a4-bat-web/1680a0c17a.
- Berendt, Bettina, and Sören Preibusch. "Better Decision Support through Exploratory Discrimination-Aware Data Mining: Foundations and Empirical Evidence." Artificial Intelligence and Law 22, no. 2 (June 2014): 175–209. https://doi.org/10.1007/s10506-013-9152-0.
- ———. "Toward Accountable Discrimination-Aware Data Mining: The Importance of Keeping the Human in the Loop—and Under the Looking Glass." Big Data 5, no. 2 (June 1, 2017): 135–52. https://doi.org/10.1089/big.2016.0055.
- Berkovsky, Shlomo, Ronnie Taib, and Dan Conway. 'How to Recommend ? User Trust Factors in Movie Recommender Systems'. In *Proceedings of the 22nd International Conference on Intelligent User Interfaces IUI '17*, 287–300, 2017. https://doi.org/10.1145/3025171.3025209.
- Bérubé, Michael. 'Equality, Freedom, and/or Justice for All: A Response to Martha Nussbaum'. *Metaphilosophy* 40, no. 3–4 (2009): 352–65. https://doi.org/10.1111/j.1467-9973.2009.01595.x.

- Bicknell, Craig. 'Homeless, but ID'd, in Seattle'. *Wired*. Accessed 23 November 2021. https://www.wired.com/1999/10/homeless-but-idd-in-seattle/.
- Binns, Reuben. 'Fairness in Machine Learning: Lessons from Political Philosophy'. In *Proceedings of Machine Learning Research*. New York City, 2018.
- ———. 'On the Apparent Conflict Between Individual and Group Fairness'. *ArXiv:1912.06883 [Cs, Stat]*, 14 December 2019. http://arxiv.org/abs/1912.06883.
- Binns, Reuben, Max Van Kleek, Michael Veale, Ulrik Lyngs, Jun Zhao, and Nigel Shadbolt. '"It's Reducing a Human Being to a Percentage"; Perceptions of Justice in Algorithmic Decisions'. *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems - CHI '18*, 2018, 1–14. https://doi.org/10.1145/3173574.3173951.
- Binns, Reuben, and Michael Veale. 'Is That Your Final Decision? Multi-Stage Profiling, Selective Effects, and Article 22 of the GDPR'. *International Data Privacy Law* 11, no. 4 (1 November 2021): 319–32. https://doi.org/10.1093/idpl/ipab020.
- Birhane, Abeba, William Isaac, Vinodkumar Prabhakaran, Mark Díaz, Madeleine Clare Elish, Iason Gabriel, and Shakir Mohamed. 'Power to the People? Opportunities and Challenges for Participatory AI', 15 September 2022. https://doi.org/10.1145/3551624.3555290.
- Boçon, Mariana Schafhauser. 'The CJEU's Judgment in CHEZ and Indirect Discrimination by Association: An Analysis of Progress in Combating Racial or Ethnic Discrimination in the European Union'. *UNIO – EU Law Journal* 3, no. 2 (12 March 2019): 20–33. https://doi.org/10.21814/unio.3.2.3.
- Bondi, Elizabeth, Lily Xu, Diana Acosta-Navas, and Jackson A. Killian. 'Envisioning Communities: A Participatory Approach Towards AI for Social Good'. In *Proceedings of the 2021 AAAI/ACM Conference on AI, Ethics, and Society*, 425–36, 2021. https://doi.org/10.1145/3461702.3462612.
- Bono, Teresa, Karen Croxson, and Adam Giles. 'Algorithmic Fairness in Credit Scoring'. *Oxford Review of Economic Policy* 37, no. 3 (23 September 2021): 585–617. https://doi.org/10.1093/oxrep/grab020.
- Borenstein, Jason, and Yvette Pearson. 'Robot Caregivers: Harbingers of Expanded Freedom for All?' *Ethics and Information Technology* 12, no. 3 (1 September 2010): 277–88. https://doi.org/10.1007/s10676-010-9236-4.
- Borgesius, Frederik J. Zuiderveen. 'Strengthening Legal Protection against Discrimination by Algorithms and Artificial Intelligence'. *The International Journal of Human Rights* 24, no. 10 (25 November 2020): 1572–93. https://doi.org/10.1080/13642987.2020.1743976.
- Boso, Luke A. 'Dignity, Inequality, and Stereotypes'. *Washington Law Review* 92 (n.d.): 66.
- Bozdag, Engin. 'Bias in Algorithmic Filtering and Personalization'. *Ethics and Information Technology* 15, no. 3 (2013): 209–27. https://doi.org/10.1007/s10676-013-9321-6.
- Bozdag, Engin, and Jeroen van den Hoven. 'Breaking the Filter Bubble: Democracy and Design'. *Ethics and Information Technology* 17, no. 4 (2015): 249–65. https://doi.org/10.1007/s10676-015-9380-y.
- Brantingham, P. Jeffrey, Matthew Valasik, and George O. Mohler. 'Does Predictive Policing Lead to Biased Arrests? Results From a Randomized Controlled Trial'. *Statistics and Public Policy* 5, no. 1 (1 January 2018): 1–6. https://doi.org/10.1080/2330443X.2018.1438940.
- Brems, Eva, and Alexandra Timmer, eds. *Stereotypes and Human Rights Law*. Intersentia, 2016. https://doi.org/10.1017/9781780685458.
- Brkan, Maja. 'The Concept of Essence of Fundamental Rights in the EU Legal Order: Peeling the Onion to Its Core'. *European Constitutional Law Review* 14, no. 2 (June 2018): 332–68. https://doi.org/10.1017/S1574019618000159.
- Büchi, Moritz, Eduard Fosch-Villaronga, Christoph Lutz, Aurelia Tamò-Larrieux, Shruthi Velidi, and Salome Viljoen. 'The Chilling Effects of Algorithmic Profiling: Mapping the Issues'. *Computer Law & Security Review* 36 (April 2020): 105367. https://doi.org/10.1016/j.clsr.2019.105367.

- Buolamwini, Joy, and Timnit Gebru. 'Gender Shades: Intersectional Accuracy Disparities in Commercial Gender Classification'. In *Conference on Fairness, Accountability and Transparency*, 77–91. PMLR, 2018. http://proceedings.mlr.press/v81/buolamwini18a.html.
- Burrell, Jenna. 'How the Machine 'thinks': Understanding Opacity in Machine Learning Algorithms'. *Big Data & Society* 3, no. 1 (2016): 2053951715622512. https://doi.org/10.1177/2053951715622512.
- Burri, Susanne, and Dagmar Schiek. 'Opportunities for Legal Responses to Intersectional Gender Discrimination?', 2009, 140.
- Buyl, Maarten, Christina Cociancig, Cristina Frattone, and Nele Roekens. 'Tackling Algorithmic Disability Discrimination in the Hiring Process: An Ethical, Legal and Technical Analysis'. In *2022 ACM Conference on Fairness, Accountability, and Transparency*, 1071–82. Seoul Republic of Korea: ACM, 2022. https://doi.org/10.1145/3531146.3533169.
- Cadwalladr, Carole, and Emma Graham-Harrison. 'How Cambridge Analytica Turned Facebook "Likes" into a Lucrative Political Tool'. *The Guardian*, 17 March 2018, sec. Technology. https://www.theguardian.com/technology/2018/mar/17/facebook-cambridge-analytica-kogan-data-algorithm.
- Cai, Han, Kan Ren, Weinan Zhang, Kleanthis Malialis, Jun Wang, Yong Yu, and Defeng Guo. 'Real-Time Bidding by Reinforcement Learning in Display Advertising'. In *Proceedings of the Tenth ACM International Conference on Web Search and Data Mining*, 661–70, 2017. https://doi.org/10.1145/3018661.3018702.
- Carastathis, Anna. 'Basements and Intersections'. *Hypatia* 28, no. 4 (2013): 698–715.
- Cartabia, M. 'The European Court of Human Rights: Judging Nondiscrimination'. *International Journal of Constitutional Law* 9, no. 3–4 (1 October 2011): 808–14. https://doi.org/10.1093/icon/mor043.
- Castelluccia, Claude, and Daniel Le Metayer. 'Understanding Algorithmic Decision-Making: Opportunities and Challenges'. European Parliament Research Service, March 2019. http://www.europarl.europa.eu/RegData/etudes/STUD/2019/624261/EPRS_STU(2019)624261_EN.pdf.
- Catherine Barnard and Bob Hepple. 'Substantive Equality'. *Cambridge Law Journal*, no. 59 (21 December 2000): 562–85.
- Microsoft Research. 'Causality and Machine Learning'. Accessed 19 May 2022. https://www.microsoft.com/en-us/research/group/causal-inference/.
- 'China "Social Credit": Beijing Sets up Huge System - BBC News'. Accessed 14 September 2019. https://www.bbc.com/news/world-asia-china-34592186.
- Citron, Daniel Keats, and Frank Pasquale. 'The Scored Society: Due Process for Automated Predictions'. *Wash. L. Rev.* 89, no. 1 (2014): 1–33.
- Citron, Danielle Keats, and Frank Pasquale. 'The Scored Society: Due Process for Automated Predictions'. *Washington Law Review* 89, no. 1 (2014): 33.
- Claassen, Rutger. 'Human Dignity in the Capability Approach'. In *The Cambridge Handbook of Human Dignity*, edited by Marcus Duwell, Jens Braarvig, Roger Brownsword, and Dietmar Mieth, 240–49. Cambridge: Cambridge University Press, 2014. https://doi.org/10.1017/CBO9780511979033.030.
- Clifford, Damian. 'Citizen-Consumers in a Personalised Galaxy: Emotion-Influenced Decision Making, a True Path to the Dark Side?' In *Future Law: Emerging Technology, Regulation and Ethics*, edited by Edina Harbinja, Lilian Edwards, and Schafer Burkhard, 79–130. Edinburgh University Press, 2020. https://www.cambridge.org/core/books/future-law/citizenconsumers-in-a-personalised-galaxy-emotioninfluenced-decision-making-a-true-path-to-the-dark-side/8EC23988B488888E1F8DC5FD15FF17D3.
- Coeckelbergh, Mark. 'Health Care, Capabilities, and AI Assistive Technologies'. *Ethical Theory and Moral Practice* 13, no. 2 (1 April 2010): 181–90. https://doi.org/10.1007/s10677-009-9186-2.

- ———. ‘Human Development or Human Enhancement? A Methodological Reflection on Capabilities and the Evaluation of Information Technologies’. *Ethics and Information Technology* 13, no. 2 (June 2011): 81–92. https://doi.org/10.1007/s10676-010-9231-9.
- ———. ‘Human Development or Human Enhancement? A Methodological Reflection on Capabilities and the Evaluation of Information Technologies’. *Ethics and Information Technology* 13, no. 2 (1 June 2011): 81–92. https://doi.org/10.1007/s10676-010-9231-9.
- Cohen, G. A. ‘On the Currency of Egalitarian Justice’. *Ethics* 99, no. 4 (1989): 906–44.
- Cohen, Julie E. *Between Truth and Power: The Legal Constructions of Informational Capitalism*. Oxford, New York: Oxford University Press, 2019.
- Collins, Hugh. ‘Discrimination, Equality and Social Inclusion’. *The Modern Law Review* 66, no. 1 (2003): 16–43. https://doi.org/10.1111/1468-2230.6601002.
- Collins, Hugh, and Tarunabh Khaitan, eds. *Foundations of Indirect Discrimination Law*. 1st ed. Oxford: Hart Publishing, 2018. https://doi.org/10.5040/9781509912551.
- ———. ‘Indirect Discrimination Law: Controversies and Critical Questions’. In *Foundations of Indirect Discrimination Law*, edited by Hugh Collins and Tarunabh Khaitan, 1st ed., 1–30. Oxford: Hart Publishing, 2018. http://www.bloomsburycollections.com/book/foundations-of-indirect-discrimination-law/ch1-indirect-discrimination-law-controversies-and-critical-questions/.
- Collins, Patricia Hill. *Fighting Words: Black Women and the Search for Justice.* Minneapolis: University of Minnesota Press, 1998. https://www.upress.umn.edu/book-division/books/fighting-words.
- ———. *Intersectionality as Critical Social Theory*. *Intersectionality as Critical Social Theory*. Duke University Press, 2019. https://www-degruyter-com.kuleuven.e-bronnen.be/document/doi/10.1515/9781478007098/html.
- ———. ‘Intersectionality’s Definitional Dilemmas’. *Annual Review of Sociology* 41, no. 1 (14 August 2015): 1–20. https://doi.org/10.1146/annurev-soc-073014-112142.
- Cook, Rebecca J., and Simone Cusack. *Gender Stereotyping : Transnational Legal Perspectives*. Philadelphia PA: University of Pennsylvania Press, 2010. https://www.worldcat.org/title/gender-stereotyping-transnational-legal-perspectives/oclc/320800041?page=citation.
- Corbyn, Zoë. ‘Hungry Judges Dispense Rough Justice’. *Nature*, 11 April 2011. https://doi.org/10.1038/news.2011.227.
- Costa, Jean-Paul. ‘Human Dignity in the Jurisprudence of the European Court of Human Rights’. In *Understanding Human Dignity*, edited by Christopher McCrudden, 0. British Academy, 2013. https://doi.org/10.5871/bacad/9780197265642.003.0022.
- Costa, Luiz. *Virtuality and Capabilities in a World of Ambient Intelligence*. Vol. 32. Law, Governance and Technology Series. Cham: Springer International Publishing, 2016. https://doi.org/10.1007/978-3-319-39198-4.
- Council of Europe. ‘Guide on Article 14 of the European Convention on Human Rights and on Article 1 of Protocol No. 12 to the Convention: Prohibition of Discrimination.’ Council of Europe, 31 December 2020. https://www.echr.coe.int/Documents/Guide_Art_14_Art_1_Protocol_12_ENG.pdf.
- Crawford, Kate. *Atlas of AI: Power, Politics, and the Planetary Costs of Artificial Intelligence*. Yale University Press, 2021. https://doi.org/doi:10.12987/9780300252392.
- ———. ‘The Hidden Biases in Big Data’. *Harvard Business Review*, 1 April 2013. https://hbr.org/2013/04/the-hidden-biases-in-big-data.
- Crawford, Kate, and Vladan Joler. ‘Anatomy of an AI System’. Anatomy of an AI System. Accessed 26 January 2022. http://www.anatomyof.ai.
- Crawford, Kate, and Trevor Paglen. ‘Excavating AI: The Politics of Images in Machine Learning Training Sets’. Excavating AI, 19 September 2019. https://www.excavating.ai.

- Crawford, Matthew T., Steven J. Sherman, and David L. Hamilton. 'Perceived Entitativity, Stereotype Formation, and the Interchangeability of Group Members.' *Journal of Personality and Social Psychology* 83, no. 5 (2002): 1076–94.
- Creel, Kathleen, and Deborah Hellman. 'The Algorithmic Leviathan: Arbitrariness, Fairness, and Opportunity in Algorithmic Decision Making Systems'. In *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency*, 816. FAccT '21. New York, NY, USA: Association for Computing Machinery, 2021. https://doi.org/10.1145/3442188.3445942.
- ———. 'The Algorithmic Leviathan: Arbitrariness, Fairness, and Opportunity in Algorithmic Decision Making Systems'. In *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency*, 816–816. Virtual Event Canada: ACM, 2021. https://doi.org/10.1145/3442188.3445942.
- ———. 'The Algorithmic Leviathan: Arbitrariness, Fairness, and Opportunity in Algorithmic Decision-Making Systems'. *Canadian Journal of Philosophy* 52, no. 1 (January 2022): 26–43. https://doi.org/10.1017/can.2022.3.
- Crenshaw, Kimberle. 'Demarginalizing the Intersection of Race and Sex: A Black Feminist Critique of Antidiscrimination Doctrine, Feminist Theory and Antiracist Politics'. *University of Chicago Legal Forum* 1989, no. 1 (1989): 31.
- ———. 'Mapping the Margins: Intersectionality, Identity Politics, and Violence against Women of Color'. *Stanford Law Review* 43, no. 6 (1991): 1241–99. https://doi.org/10.2307/1229039.
- Croon-Gestefeld, Johanna. *Reconceptualising European Equality Law: A Comparative Institutional Analysis*. 1st ed. Modern Studies in European Law. Oxford: Hart Publishing, 2017. https://doi.org/10.5040/9781509909674.
- Cudd, Ann E. *Oppression: The Fundamental Injustice of Social Institutions*. *Analyzing Oppression*. Oxford University Press. Accessed 7 June 2021. http://oxford.universitypressscholarship.com/view/10.1093/0195187431.001.0001/acprof-9780195187434-chapter-1.
- ———. 'Psychological Harms of Oppression'. In *Analyzing Oppression*. New York: Oxford University Press, 2006. https://doi.org/10.1093/0195187431.003.0006.
- Custers, Bart, Toon Calders, Bart Schermer, and Tal Zarsky, eds. *Discrimination and Privacy in the Information Society: Data Mining and Profiling in Large Databases*. Studies in Applied Philosophy, Epistemology and Rational Ethics. Berlin Heidelberg: Springer-Verlag, 2013. https://www.springer.com/gp/book/9783642304866.
- Danaher, John. 'Freedom in an Age of Algocracy'. The Oxford Handbook of Philosophy of Technology, 10 May 2022. https://doi.org/10.1093/oxfordhb/9780190851187.013.16.
- ———. 'Philosophical Disquisitions: Algorithmic Micro-Domination: Living with Algocracy'. *Philosophical Disquisitions* (blog), 18 June 2018. https://philosophicaldisquisitions.blogspot.com/2018/06/algorithmic-micro-domination-living.html.
- Danisi, C. 'How Far Can the European Court of Human Rights Go in the Fight against Discrimination? Defining New Standards in Its Nondiscrimination Jurisprudence'. *International Journal of Constitutional Law* 9, no. 3–4 (1 October 2011): 793–807. https://doi.org/10.1093/icon/mor044.
- Danziger, Shai, Jonathan Levav, and Liora Avnaim-Pesso. 'Extraneous Factors in Judicial Decisions'. *Proceedings of the National Academy of Sciences* 108, no. 17 (26 April 2011): 6889–92. https://doi.org/10.1073/pnas.1018033108.
- Datatilsynet (Norwegian Data Protection Authority). 'Advance Notification of Order to Rectify Unfairly Processed and Incorrect Personal Data - International Baccalaureate Organization', 7 August 2020. https://www.datatilsynet.no/contentassets/04df776f85f64562945f1d261b4add1b/advance-notification-of-order-to-rectify-unfairly-processed-and-incorrect-personal-data.pdf.

- Dawson, Mark. *The Governance of EU Fundamental Rights*. Cambridge: Cambridge University Press, 2017. https://doi.org/10.1017/9781107707436.
- Dawson, Mark, Orla Lynskey, and Elise Muir. 'What Is the Added Value of the Concept of the "Essence" of EU Fundamental Rights?' *German Law Journal* 20, no. 6 (September 2019): 763–78. https://doi.org/10.1017/glj.2019.61.
- De Beco, Gauthier. 'Harnessing the Full Potential of Intersectionality Theory in International Human Rights Law : Lessons from Disabled Children's Right to Education'. In *Intersectionality and Human Rights Law*, edited by Shreya Atrey and Peter Dunne, 1st ed., 39–58. Oxford: Hart Publishing, 2020. http://www.bloomsburycollections.com/book/intersectionality-and-human-rights-law/ch2-harnessing-the-full-potential-of-intersectionality-theory-in-international-human-rights-law/.
- De La Corte-Rodríguez, Miguel. 'The First Revision of the Hofmann Case Law on Maternity Leave and Discrimination Against Fathers: Care-Giving as the Pivot?' *Industrial Law Journal* 50, no. 2 (1 June 2021): 306–23. https://doi.org/10.1093/indlaw/dwab008.
- De Schutter, Olivier. 'Three Models of Equality and European Anti-Discrimination Law'. *Northern Ireland Legal Quarterly* 57, no. 1 (2006): 1–56.
- De Vos, Marc. 'Substantive Formal Equality in EU Non-Discrimination Law'. In *The European Union as Protector and Promoter of Equality*, edited by Thomas Giegerich, 245–67. European Union and Its Neighbours in a Globalized World. Cham: Springer International Publishing, 2020. https://doi.org/10.1007/978-3-030-43764-0_13.
- ———. 'The European Court of Justice and the March towards Substantive Equality in European Union Anti-Discrimination Law'. *International Journal of Discrimination and the Law* 20, no. 1 (1 March 2020): 62–87. https://doi.org/10.1177/1358229120927947.
- Merriam-Webster. 'Definition of Stereotype', 2022. https://www.merriam-webster.com/dictionary/stereotype.
- Delacroix, Sylvie. 'Beware of "Algorithmic Regulation"'. SSRN Scholarly Paper. Rochester, NY: Social Science Research Network, 1 February 2019. https://doi.org/10.2139/ssrn.3327191.
- Delacroix, Sylvie, and Michael Veale. 'Smart Technologies and Our Sense of Self: Going beyond Epistemic Counter-Profiling'. *Life and the Law in the Era of Data-Driven Agency*, 31 January 2020. https://www.elgaronline.com/view/edcoll/9781788971997/9781788971997.00011.xml.
- Dencik, Lina, and Javier Sanchez-Monedero. 'Data Justice'. *Internet Policy Review* 11, no. 1 (14 January 2022). https://policyreview.info/articles/analysis/data-justice.
- Deneulin, Severine. 'Recovering Nussbaum's Aristotelian Roots'. *International Journal of Social Economics* 40, no. 7 (2013): 624–32. https://doi.org/10.1108/IJSE-2012-0127.
- Denton, Emily, Alex Hanna, Razvan Amironesei, Andrew Smart, Hilary Nicole, and Morgan Klaus Scheuerman. 'Bringing the People Back In: Contesting Benchmark Machine Learning Datasets'. *ArXiv:2007.07399 [Cs]*, 14 July 2020. http://arxiv.org/abs/2007.07399.
- D'Hooghe, Eugenie. 'Ook in Brussel minder kans bij sollicitaties voor mensen met migratieroots'. De Morgen, 20 March 2021. https://www.demorgen.be/gs-bd34de11.
- Diakopoulos, Nicholas. 'Accountability in Algorithmic Decision Making'. *Commun. ACM* 59, no. 2 (January 2016): 56–62. https://doi.org/10.1145/2844110.
- D'Ignazio, Catherine, and Lauren F. Klein. *Data Feminism*. Strong Ideas Series. Cambridge, Massachusetts: The MIT Press, 2020. https://mitpress.mit.edu/books/data-feminism.
- Directorate-General for Justice and Consumers (European Commission), European network of legal experts in gender equality and non-discrimination, Colm O'Cinneide, and Kimberly Liu. *The Ongoing Evolution of the Case-Law of the Court of Justice of the European Union on Directives 2000/43/EC and 2000/78/EC: A Legal Analysis of*

the Situation in EU Member States. LU: Publications Office of the European Union, 2019. https://data.europa.eu/doi/10.2838/307887.

- Donegan, Cara. 'Thinly Veiled Discrimination : Muslim Women, Intersectionality and the Hybrid Solution of Reasonable Accommodation and Proactive Measures'. *European Journal of Legal Studies*, no. 2 (30 November 2020): 143–79. https://doi.org/10.2924/EJLS.2019.032.
- Dorloff, Axel, and Daniel Satra. 'China: Auf dem Weg zur totalen Überwachung'. *tagesschau.de*, 24 March 2019. https://www.tagesschau.de/ausland/ueberwachung-china-101.html.
- Dou, Zi-Yi, Anamika Barman-Adhikari, Fei Fang, and Amulya Yadav. 'Harnessing Social Media to Identify Homeless Youth At-Risk of Substance Use'. *Proceedings of the AAAI Conference on Artificial Intelligence* 35, no. 17 (18 May 2021): 14748–56.
- Durán, Juan Manuel, and Karin Rolanda Jongsma. 'Who Is Afraid of Black Box Algorithms? On the Epistemological and Ethical Basis of Trust in Medical AI'. *Journal of Medical Ethics* 47, no. 5 (1 May 2021): 329–35. https://doi.org/10.1136/medethics-2020-106820.
- Dwork, Cynthia, Moritz Hardt, Toniann Pitassi, Omer Reingold, and Rich Zemel. 'Fairness Through Awareness'. *ArXiv:1104.3913 [Cs]*, 19 April 2011. http://arxiv.org/abs/1104.3913.
- Dworkin, Ronald. 'Comment on Narveson: In Defense of Equality'. *Social Philosophy and Policy* 1, no. 1 (1983): 24.
- ———. *Sovereign Virtue: The Theory and Practice of Equality*. 2nd ed. Cambridge, Massachusetts ; London, England: Harvard University Press, 2000.
- ———. *Taking Rights Seriously*. Cambridge: Harvard University Press, 1977.
- ———. 'What Is Equality? Part 2: Equality of Resources'. *Philosophy & Public Affairs* 10, no. 4 (1981): 283–345.
- Eder, Niklas. 'Privacy, Non-Discrimination and Equal Treatment: Developing a Fundamental Rights Response to Behavioural Profiling'. In *Algorithmic Governance and Governance of Algorithms: Legal and Ethical Challenges*, edited by Martin Ebers and Marta Cantero Gamito, 23–48. Data Science, Machine Intelligence, and Law. Cham: Springer International Publishing, 2021. https://doi.org/10.1007/978-3-030-50559-2_2.
- Edmonds, David. *Caste Wars: A Philosophy of Discrimination*. London; New York: Routledge, 2006. http://site.ebrary.com/id/10272708.
- Edwards, Lilian, and Michael Veale. 'Slave to the Algorithm? Why a "right to an Explanation" Is Probably Not the Remedy You Are Looking For.' *Duke Law & Technology Review* 16, no. 1 (23 May 2017): 18–83. https://doi.org/Edwards, Lilian and Veale, Michael, Slave to the Algorithm? Why a 'Right to an Explanation' Is Probably Not the Remedy You Are Looking For (May 23, 2017). 16 Duke Law & Technology Review 18 (2017). Available at SSRN: https://ssrn.com/abstract=2972855 or http://dx.doi.org/10.2139/ssrn.2972855.
- Eidelson, Benjamin. 'Treating People as Individuals'. In *Philosophical Foundations of Discrimination Law*. Oxford: Oxford University Press, 2013. https://doi.org/10.1093/acprof:oso/9780199664313.003.0011.
- Eklund, Hanna, and Claire Kilpatrick. 'Article 21 EU Charter of Fundamental Rights'. In *The EU Charter of Fundamental Rights: A Commentary*, edited by S Peers, J Hervey, J Kenner, and A Ward, 33. Hart Publishing, 2021.
- Ellis, Evelyn, and Philippa Watson. 'Key Concepts in EU Anti-Discrimination Law'. In *EU Anti-Discrimination Law*, 2nd ed. Oxford: Oxford University Press, 2012. https://doi.org/10.1093/acprof:oso/9780199698462.003.0004.
- Elster, Jon. *Sour Grapes: Studies in the Subversion of Rationality*. Cambridge Philosophy Classics. Cambridge: Cambridge University Press, 2016. https://doi.org/10.1017/CBO9781316494172.
- Ensign, Danielle, Sorelle A. Friedler, Scott Neville, Carlos Scheidegger, and Suresh Venkatasubramanian. 'Runaway Feedback Loops in Predictive Policing'. *ArXiv:1706.09847 [Cs, Stat]*, 29 June 2017. http://arxiv.org/abs/1706.09847.
- Equivant. 'Practicioner's Guide to COMPAS Core'. Equivant (Northpointe Inc.), 4 April 2019. https://www.equivant.com/wp-content/uploads/Practitioners-Guide-to-COMPAS-Core-040419.pdf.

- Commissioner for Human Rights. 'Ethnic Profiling: A Persisting Practice in Europe'. Accessed 21 September 2022. https://www.coe.int/en/web/commissioner/blog/-/asset_publisher/xZ32OPEoxOkq/content/ethnic-profiling-a-persisting-practice-in-europe.
- Eubanks, Virginia. *Automating Inequality: How High-Tech Tools Profile, Police and Punish the Poor*. Sint Martin's Press: Macmillan Publishers, 2018.
- ———. 'The Digital Poorhouse: Permitless Carry and the New Gun-Rights Extremism'. *Harper's Magazine*, 1 January 2018. https://harpers.org/archive/2018/01/the-digital-poorhouse/.
- Europäische Union, ed. *Understanding and Preventing Discriminatory Ethnic Profiling: A Guide ; towards More Effective Policing*. Luxembourg: Office for Official Publ. of the Europ. Union, 2010.
- Europäische Union, Europarat, and Europäischer Gerichtshof für Menschenrechte, eds. *Handbook on European Non-Discrimination Law*. 2018 edition. Handbook / FRA, European Union Agency for Fundamental Rights. Luxembourg: Publications Office of the European Union, 2018.
- Favaretto, Maddalena, Eva De Clercq, and Bernice Simone Elger. 'Big Data and Discrimination: Perils, Promises and Solutions. A Systematic Review'. *Journal of Big Data* 6, no. 1 (5 February 2019): 12. https://doi.org/10.1186/s40537-019-0177-4.
- Feldman, Fred, and Brad Skow. 'Desert'. In *The Stanford Encyclopedia of Philosophy*, edited by Edward N. Zalta, Winter 2020. Metaphysics Research Lab, Stanford University, 2020. https://plato.stanford.edu/archives/win2020/entries/desert/.
- Ferguson, Andrew Guthrie. 'Predictive Policing and Reasonable Suspicion'. *EMORY LAW JOURNAL* 62 (n.d.): 68.
- Ferri, Delia. 'The Unorthodox Relationship between the EU Charter of Fundamental Rights, the UN Convention on the Rights of Persons with Disabilities and Secondary Rights in the Court of Justice Case Law on Disability Discrimination'. *European Constitutional Law Review* 16, no. 2 (June 2020): 275–305. https://doi.org/10.1017/S1574019620000164.
- Fink, Katherine. 'Opening the Government's Black Boxes: Freedom of Information and Algorithmic Accountability'. *Information, Communication & Society* 21, no. 10 (3 October 2018): 1453–71. https://doi.org/10.1080/1369118X.2017.1330418.
- Fisher, Angelina, and Thomas Streinz. 'Confronting Data Inequality'. SSRN Scholarly Paper. Rochester, NY: Social Science Research Network, 1 April 2021. https://doi.org/10.2139/ssrn.3825724.
- Flaxman, Seth, Sharad Goel, and Justin M. Rao. 'Filter Bubbles, Echo Chambers, and Online News Consumption'. *Public Opinion Quarterly* 80, no. S1 (2016): 298–320. https://doi.org/10.1093/poq/nfw006.
- Fletcher. 'Gender Inequality in the Gig Economy'. *Autonomy* (blog), June 2021. https://autonomy.work/portfolio/fletcher-gig-economy/.
- Fosch-Villaronga, E., A. Poulsen, R. A. Søraa, and B. H. M. Custers. 'A Little Bird Told Me Your Gender: Gender Inferences in Social Media'. *Information Processing & Management* 58, no. 3 (1 May 2021): 102541. https://doi.org/10.1016/j.ipm.2021.102541.
- Fourie, Carina, Fabian Schuppert, and Ivo Walliman-Helmer, eds. *Social Equality: On What It Means to Be Equals*. New York: Oxford University Press, 2015.
- ———. 'The Nature and Distinctiveness of Social Equality: An Introduction'. In *Social Equality: On What It Means to Be Equals*, edited by Carina Fourie, Fabian Schuppert, and Ivo Walliman-Helmer, 1–20. New York: Oxford University Press, 2015.
- Foxman, Daniel, and Gregory Bateson. 'Steps to an Ecology of Mind'. *The Western Political Quarterly* 26, no. 2 (June 1973): 345. https://doi.org/10.2307/446833.
- BBC News. 'Frances Haugen Says Facebook Is "Making Hate Worse"', 25 October 2021, sec. Technology. https://www.bbc.com/news/technology-59038506.

- Frank, Brenda Medina, Thomas. 'TSA Agents Say They're Not Discriminating Against Black Women, But Their Body Scanners Might Be'. ProPublica. Accessed 3 August 2022. https://www.propublica.org/article/tsa-not-discriminating-against-black-women-but-their-body-scanners-might-be.
- Fraser, Nancy. 'Recognition without Ethics?' *Theory, Culture & Society* 18, no. 2–3 (1 June 2001): 21–42. https://doi.org/10.1177/02632760122051760.
- ———. 'Social Justice in the Age of Identity Politics. Redistribution, Recognition, Participation.', WZB Discussion Paper, no. FS I 98-108 (1998): 29.
- Fraser, Nancy, and Axel Honneth. *Redistribution Or Recognition?: A Political-Philosophical Exchange*. Verso, 2003.
- Fredman FBA, Sandra. *Discrimination Law*. 2nd ed. Oxford ; New York: Oxford University Press, 2011.
- Fredman, Sandra. 'Direct and Indirect Discrimination: Is There Still a Divide?' In *Foundations of Indirect Discrimination Law*, edited by Hugh Collins and Tarunabh Khaitan, 1st ed., 31–56. Oxford: Hart Publishing, 2018. http://www.bloomsburycollections.com/book/foundations-of-indirect-discrimination-law/ch2-direct-and-indirect-discrimination-is-there-still-a-divide/.
- ———. 'Emerging from the Shadows: Substantive Equality and Article 14 of the European Convention on Human Rights'. *Human Rights Law Review* 16, no. 2 (16 April 2016): 273–301. https://doi.org/10.1093/hrlr/ngw001.
- ———. 'Equality: A New Generation'. *Industrial Law Journal* 30 (2001): 145.
- ———. 'Providing Equality: Substantive Equality and the Positive Duty to Provide'. *South African Journal on Human Rights* 21 (2005): 163.
- ———. 'Redistribution and Recognition : Reconciling Inequalities'. *South African Journal on Human Rights* 23, no. 2 (January 2007): 214–34. https://doi.org/10.1080/19962126.2007.11864923.
- ———. 'Substantive Equality Revisited'. *International Journal of Constitutional Law* 14, no. 3 (July 2016): 712–38. https://doi.org/10.1093/icon/mow043.
- OHCHR. 'Freedom of Thought Increasingly Violated Worldwide, UN Expert Warns'. Accessed 30 September 2022. https://www.ohchr.org/en/press-releases/2021/10/freedom-thought-increasingly-violated-worldwide-un-expert-warns.
- Freeman, Samuel. 'Utilitarianism, Deontology, and the Priority of Right'. *Philosophy & Public Affairs* 23, no. 4 (1994): 313–49.
- Friedman, Batya, and Helen Nissenbaum. 'Bias in Computer Systems'. *ACM Transactions on Information Systems* 14, no. 3 (1 July 1996): 330–47. https://doi.org/10.1145/230538.230561.
- Frostell, Katarina. 'Gender-Based Violence Triggers Differential Treatment in Housing Benefit Case'. *Strasbourg Observers* (blog), 2 December 2019. https://strasbourg.weichie.dev/2019/12/02/gender-based-violence-triggers-differential-treatment-in-housing-benefit-case/.
- Frye, Marilyn. *The Politics of Reality: Essays in Feminist Theory*. The Crossing Press, 1983.
- Fulcher, John. 'Computational Intelligence: An Introduction'. *Computational Intelligence*, n.d., 75.
- Fuster, Gloria González, Serge Gutwirth, and Erika Ellyne. 'Profiling in the European Union':, June 2010, 12.
- Gabriel, Iason. 'Towards a Theory of Justice for Artificial Intelligence'. *Daedalus* 151, no. 2 (forthcoming spring 2022): 12.
- Gage, Deborah. 'Big Data Uncovers Some Weird Correlations'. *Wall Street Journal*, 23 March 2014, sec. Special. https://online.wsj.com/article/SB10001424052702303369904579423132072969654.html.
- Gajane, Pratik, and Mykola Pechenizkiy. 'On Formalizing Fairness in Prediction with Machine Learning'. *ArXiv:1710.03184 [Cs, Stat]*, 9 October 2017. http://arxiv.org/abs/1710.03184.
- Gallie, W. B. 'IX.—Essentially Contested Concepts'. *Proceedings of the Aristotelian Society* 56, no. 1 (1 June 1956): 167–98. https://doi.org/10.1093/aristotelian/56.1.167.
- Galston, William Arthur. *Justice and the Human Good*. University of Chicago Press, 1980.

- Ganty, Sarah. 'Poverty as Misrecognition: What Role for Anti-Discrimination Law in Europe?' *Human Rights Law Review*, no. 00 (2021): 46.
- Ganty, Sarah, and Jua Carlos Benito Sanchez. 'Expanding the List of Protected Grounds within Anti-Discrimination Law in the EU: AN Equinet Report'. Equinet, 2021.
- Garrau, Marie, and Cécile Laborde. 'Relational Equality, Non-Domination, and Vulnerability'. In *Social Equality: On What It Means to Be Equals*, edited by Carina Fourie, Fabian Schuppert, and Ivo Walliman-Helmer, 45–64. New York: Oxford University Press, 2015.
- Gebru, Timnit, Jamie Morgenstern, Briana Vecchione, Jennifer Wortman Vaughan, Hanna Wallach, Hal Daumé III, and Kate Crawford. 'Datasheets for Datasets'. *ArXiv:1803.09010 [Cs]*, 19 March 2020. http://arxiv.org/abs/1803.09010.
- Gentzel, Michael. 'Biased Face Recognition Technology Used by Government: A Problem for Liberal Democracy'. *Philosophy & Technology* 34, no. 4 (2021): 1639–63. https://doi.org/10.1007/s13347-021-00478-z.
- Gerards, J. 'The Discrimination Grounds of Article 14 of the European Convention on Human Rights'. *Human Rights Law Review* 13, no. 1 (1 March 2013): 99–124. https://doi.org/10.1093/hrlr/ngs044.
- Gerards, Janneke. *Judicial Review in Equal Treatment Cases*. Brill Nijhoff, 2005. http://brill.com/view/title/11318.
- ———. 'Margin of Appreciation and Incrementalism in the Case Law of the European Court of Human Rights'. *Human Rights Law Review* 18, no. 3 (1 September 2018): 495–515. https://doi.org/10.1093/hrlr/ngy017.
- ———. 'The Margin of Appreciation Doctrine, the Very Weighty Reasons Test and Grounds of Discrimination'. SSRN Scholarly Paper. Rochester, NY: Social Science Research Network, 1 May 2017. https://doi.org/10.2139/ssrn.2875230.
- Gerards, Janneke H. 'Intensity of Judicial Review in Equal Treatment Cases'. *Netherlands International Law Review* 51, no. 2 (September 2004): 135–83.
- Gerards, Janneke, Raphaele Xenidis, Directorate-General for Justice and Consumers (European Commission), and European network of legal experts in gender equality and non-discrimination. *Algorithmic Discrimination in Europe: Challenges and Opportunities for Gender Equality and Non Discrimination Law*. LU: Publications Office of the European Union, 2021. https://data.europa.eu/doi/10.2838/544956.
- Gerards, Janneke, and Frederik Zuiderveen Borgesius. 'Protected Grounds and the System of Non-Discrimination Law in the Context of Algorithmic Decision-Making and Artificial Intelligence'. SSRN Scholarly Paper. Rochester, NY: Social Science Research Network, 2 November 2020. https://papers.ssrn.com/abstract=3723873.
- ———. 'Protected Grounds and the System of Non-Discrimination Law in the Context of Algorithmic Decision-Making and Artificial Intelligence'. *Colorado Technology Law Journal* 20 (June 2022): 56.
- Goddard, Kate, Abdul Roudsari, and Jeremy C Wyatt. 'Automation Bias: A Systematic Review of Frequency, Effect Mediators, and Mitigators'. *Journal of the American Medical Informatics Association : JAMIA* 19, no. 1 (2012): 121–27. https://doi.org/10.1136/amiajnl-2011-000089.
- Goldberg, Suzanne. 'Discrimination by Comparison'. *Yale L.J.* 120 (1 January 2011): 728.
- Goldberg, Suzanne B. 'Discrimination by Comparison'. *The Yale Law Journal*, 2011, 85.
- Gomberg, Paul. 'Dilemmas of Rawlsian Opportunity'. *Canadian Journal of Philosophy* 40, no. 1 (2010): 1–24.
- Goodman, Bryce, and Seth Flaxman. 'EU Regulations on Algorithmic Decision-Making and a "Right to Explanation"'. *2016 ICML Workshop on Human Interpretability in Machine Learning (WHI 2016)*, 2016, 1–9.
- Gosepath, Stefan. 'Equality'. In *The Stanford Encyclopedia of Philosophy*, edited by Edward N. Zalta, Spring 2011. Metaphysics Research Lab, Stanford University, 2011. https://plato.stanford.edu/archives/spr2011/entries/equality/.
- Gould, Carol C. *Rethinking Democracy: Freedom and Social Cooperation in Politics, Economy, and Society*. Cambridge University press, 1988.

- Grace, Jamie, and Roxanne Bamford. '"AI Theory of Justice": Using Rawlsian Approaches to Better Legislate on Machine Learning in Government'. *SSRN Electronic Journal*, 2020. https://doi.org/10.2139/ssrn.3588256.
- Gräf, Eike. 'When Automated Profiling Threatens Our Freedom': *European Data Protection Law Review* 3, no. 4 (2017): 441–51. https://doi.org/10.21552/edpl/2017/4/6.
- Greenwald, Glenn, Ewen MacAskill, Laura Poitras, Spencer Ackerman, and Dominic Rushe. 'Microsoft Handed the NSA Access to Encrypted Messages'. *The Guardian*, 12 July 2013, sec. US news. https://www.theguardian.com/world/2013/jul/11/microsoft-nsa-collaboration-user-data.
- Grosan, Crina, and Ajith Abraham. *Intelligent Systems*. Vol. 17. Intelligent Systems Reference Library. Berlin, Heidelberg: Springer Berlin Heidelberg, 2011. https://doi.org/10.1007/978-3-642-21004-4.
- Guidotti, Riccardo, Anna Monreale, Salvatore Ruggieri, Franco Turini, Fosca Giannotti, and Dino Pedreschi. 'A Survey of Methods for Explaining Black Box Models'. *ACM Comput. Surv.* 51, no. 5 (August 2018): 93:1-93:42. https://doi.org/10.1145/3236009.
- Gurol, Senem. 'Challenging Gender Stereotyping before the ECtHR: Case of Carvalho Pinto v. Portugal'. *EJIL: Talk!* (blog), 21 September 2017. https://www.ejiltalk.org/challenging-gender-stereotyping-before-the-ecthr-case-of-carvalho-pinto-v-portugal/.
- Gutwirth, Serge, and Paul De Hert. 'Regulating Profiling in a Democratic Constitutional State'. In *Profiling the European Citizen: Cross-Disciplinary Perspectives*, edited by Mireille Hildebrandt and Serge Gutwirth, 271–302. Dordrecht: Springer Netherlands, 2008. https://doi.org/10.1007/978-1-4020-6914-7_14.
- Guyer, Paul. 'Principles of Justice, Primary Goods and Categories of Right: Rawls and Kant'. *Kantian Review* 23, no. 4 (December 2018): 581–613. https://doi.org/10.1017/S1369415418000419.
- Hacker, Philipp. 'Manipulation by Algorithms. Exploring the Triangle of Unfair Commercial Practice, Data Protection, and Privacy Law'. *European Law Journal* n/a, no. n/a. Accessed 10 May 2022. https://doi.org/10.1111/eulj.12389.
- Hacking, Ian. 'Making up People'. In *Reconstructing Individualism: Autonomy, Individuality, and the Self in Western Thought.*, edited by T.C. Heller, M. Sosna, and D.E. Wellbery, 222–36. Stanford: Stanford University Press, 1986.
- Hampton, Jean. *Political Philosophy*. Vol. 108. Westview Press, 1997.
- Han, Hye Jung. 'An Algorithm Shouldn't Decide a Student's Future'. POLITICO, 13 August 2020. https://www.politico.eu/article/an-algorithm-shouldnt-decide-students-future-coronavirus-international-baccalaureate/.
- Harambam, Jaron, Natali Helberger, and Joris van Hoboken. 'Democratizing Algorithmic News Recommenders: How to Materialize Voice in a Technologically Saturated Media Ecosystem'. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences* 376, no. 2133 (28 November 2018): 20180088. https://doi.org/10.1098/rsta.2018.0088.
- Hart, H. L. A. *Rawls on Liberty and Its Priority*. *Essays in Jurisprudence and Philosophy*. Oxford University Press. Accessed 20 August 2020. http://oxford.universitypressscholarship.com/view/10.1093/acprof:oso/9780198253884.001.0001/acprof-9780198253884-chapter-11.
- Havelková, Barbara. 'Judicial Scepticism of Discrimination at the ECtHR'. In *Foundations of Indirect Discrimination Law*, edited by Hugh Collins and Tarunabh Khaitan, 1st ed., 83–104. Oxford: Hart Publishing, 2018. http://www.bloomsburycollections.com/book/foundations-of-indirect-discrimination-law/ch4-judicial-scepticism-of-discrimination-at-the-ecthr/.
- Hayes, Paul, Ibo van de Poel, and Marc Steen. 'Algorithms and Values in Justice and Security'. *AI & SOCIETY* 35, no. 3 (September 2020): 533–55. https://doi.org/10.1007/s00146-019-00932-9.

- Hayes, Story by Allison Hope, Illustrations by India. 'Traveling as a Trans Person: It's Complicated'. CNN. Accessed 29 April 2022. https://www.cnn.com/travel/article/transgender-travel-complications-tips/index.html.
- Helberger, Natali, Kari Karppinen, and Lucia D'Acunto. 'Exposure Diversity as a Design Principle for Recommender Systems'. *Information, Communication & Society* 21, no. 2 (February 2018): 191–207. https://doi.org/10.1080/1369118X.2016.1271900.
- Hellman, Deborah. 'Discrimination and Social Meaning'. In *The Routledge Handbook of the Ethics of Discrimination*. Routledge, 2017.
- Hellman, Deborah, and Sophia Moreau, eds. *Philosophical Foundations of Discrimination Law*. Philosophical Foundations of Law. Oxford: Oxford University Press, 2013. https://doi.org/10.1093/acprof:oso/9780199664313.001.0001.
- Hepple, Bob. 'Discrimination and Equality of Opportunity—Northern Irish Lessons'. *Oxford Journal of Legal Studies* 10, no. 3 (1990): 408–21. https://doi.org/10.1093/ojls/10.3.408.
- Heri, Corina. 'Between a Rock and a Hard Place: The Court's Difficult Choice in Khamtokhu and Aksenchik v. Russia'. *Strasbourg Observers* (blog), 17 March 2017. https://strasbourg.weichie.dev/2017/03/17/between-a-rock-and-a-hard-place-the-courts-difficult-choice-in-khamtokhu-and-aksenchik-v-russia/.
- Heringa, Aalt Willem. 'Standards of Review for Discrimination. The Scope of Review by the Courts.' In *Non-Discrimination Law: Comporative Perspectives*, edited by Titia Loenen and Peter R. Rodrigues. The Hague: Kluwer Law International, 1999.
- Hern, Alex. 'Ofqual's A-Level Algorithm: Why Did It Fail to Make the Grade?' the Guardian, 21 August 2020. http://www.theguardian.com/education/2020/aug/21/ofqual-exams-algorithm-why-did-it-fail-make-grade-a-levels.
- High Level Expert Group on AI. 'Draft Ethics Guidelines for Trustworthy AI'. Text. Shaping Europe's digital future - European Commission, 18 December 2018. https://ec.europa.eu/digital-single-market/en/news/draft-ethics-guidelines-trustworthy-ai.
- ———. 'The Assessment List for Trustworthy Artificial Intelligence (ALTAI) for Self-Assessment', 17 July 2020.
- High-Level Expert Group on Artificial Intelligence. 'Ethics Guidelines for Trustworthy AI', April 2019. https://ec.europa.eu/digital-single-market/en/news/ethics-guidelines-trustworthy-ai.
- Hildebrandt, M. 'Discrimination, Data-Driven AI Systems and Practical Reason'. *European Data Protection Law Review* 7, no. 3 (2021): 358–66. https://doi.org/10.21552/edpl/2021/3/4.
- ———. 'Discrimination, Data-Driven AI Systems and Practical Reason'. *European Data Protection Law Review* 7, no. 3 (2021): 358–66. https://doi.org/10.21552/edpl/2021/3/4.
- Hildebrandt, Mireille. 'Algorithmic Regulation and the Rule of Law'. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences* 376, no. 2128 (13 September 2018): 20170355. https://doi.org/10.1098/rsta.2017.0355.
- ———. 'Defining Profiling: A New Type of Knowledge?' In *Profiling the European Citizen: Cross-Disciplinary Perspectives*, edited by Mireille Hildebrandt and Serge Gutwirth, 17–45. Dordrecht: Springer Netherlands, 2008. https://doi.org/10.1007/978-1-4020-6914-7_2.
- Hildebrandt, Mireille, and Serge Gutwirth, eds. *Profiling the European Citizen: Cross-Disciplinary Perspectives*. New York: Springer, 2008.
- Hildebrandt, Mireille, and Bert-Jaap Koops. 'The Challenges of Ambient Law and Legal Protection in the Profiling Era: The Challenges of Ambient Law and Legal Protection in the Profiling Era'. *Modern Law Review* 73, no. 3 (May 2010): 428–60. https://doi.org/10.1111/j.1468-2230.2010.00806.x.

- Hill, Kashmir. 'Another Arrest, and Jail Time, Due to a Bad Facial Recognition Match'. *The New York Times*, 29 December 2020, sec. Technology. https://www.nytimes.com/2020/12/29/technology/facial-recognition-misidentify-jail.html.
- Hoffmann, Anna Lauren. 'Rawls, Information Technology, and the Sociotechnical Bases of Self-Respect'. The Oxford Handbook of Philosophy of Technology, 10 May 2022. https://doi.org/10.1093/oxfordhb/9780190851187.013.15.
- ———. 'Where Fairness Fails: Data, Algorithms, and the Limits of Antidiscrimination Discourse'. *Information, Communication & Society* 22, no. 7 (7 June 2019): 900–915. https://doi.org/10.1080/1369118X.2019.1573912.
- Holmes, Elisa. 'Anti-Discrimination Rights without Equality'. *The Modern Law Review* 68, no. 2 (2005): 175–94.
- Holzleithner, Elisabeth. 'Mainstreaming Equality: Dis/Entangling Grounds of Discrimination Symposium: Antidiscrimination Law in the European Union'. *Transnational Law & Contemporary Problems* 14, no. 3 (2005 2004): 927–58.
- Honneth, Axel. 'Recognition and Justice: Outline of a Plural Theory of Justice'. *Acta Sociologica* 47, no. 4 (December 2004): 351–64. https://doi.org/10.1177/0001699304048668.
- ———. *The Struggle for Recognition: The Moral Grammar of Social Conflicts*. Translated by Joel Anderson. Studies in Contemporary German Social Thought. Cambridge, MA, USA: MIT Press, 1996.
- Access Now. 'How AI Systems Undermine LGBTQ Identity', 6 April 2021. https://www.accessnow.org/how-ai-systems-undermine-lgbtq-identity/.
- VentureBeat. 'How Reinforcement Learning Chooses the Ads You See', 23 February 2021. https://venturebeat.com/2021/02/23/how-reinforcement-learning-chooses-the-ads-you-see/.
- Howard, Erica. 'Headscarves and the CJEU: Protecting Fundamental Rights and Pandering to Prejudice: The CJEU Does Both'. *Maastricht Journal of European and Comparative Law*, 28 February 2022, 1023263X221080557. https://doi.org/10.1177/1023263X221080557.
- Hoye, J. Matthew, and Jeffrey Monaghan. 'Surveillance, Freedom and the Republic'. *European Journal of Political Theory* 17, no. 3 (July 2018): 343–63. https://doi.org/10.1177/1474885115608783.
- Huisman, Charlotte. 'SyRI, het fraudesysteem van de overheid, faalt: nog niet één fraudegeval opgespoord'. de Volkskrant, 27 June 2019. https://www.volkskrant.nl/gs-b789bc3a.
- Hunt, Elle. 'Tay, Microsoft's AI Chatbot, Gets a Crash Course in Racism from Twitter'. *The Guardian*, 24 March 2016, sec. Technology. https://www.theguardian.com/technology/2016/mar/24/tay-microsofts-ai-chatbot-gets-a-crash-course-in-racism-from-twitter.
- Hurley, Mikella, and Julius Adebayo. 'Credit Scoring in the Era of Big Data'. *Yale J.L. & Tech* 18 (2016): 148–216.
- Huszár, Ferenc, Sofia Ira Ktena, Conor O'Brien, Luca Belli, Andrew Schlaikjer, and Moritz Hardt. 'Algorithmic Amplification of Politics on Twitter', 2021, 27.
- AlgorithmWatch. 'In French Daycare, Algorithms Attempt to Fight Cronyism'. Accessed 26 April 2022. https://algorithmwatch.org/en/algorithms-to-fight-cronyism-in-french-daycare/.
- International Baccalaureate. 'IB Assessment for May 2020 Examination Session during COVID-19 Pandemic', 2020. https://www.ibo.org/globalassets/news-assets/coronavirus/assessment-model-letter-may-2020-en.pdf.
- Ippolito, Francesca, and Sara Iglesias Sánchez. 'Introduction'. In *Protecting Vulnerable Groups : The European Human Rights Framework*, edited by Francesca Ippolito and Sara Iglesias Sánchez, 1st ed., 1–20. Modern Studies in European Law. London: Hart Publishing, 2015. http://www.bloomsburycollections.com/book///.
- Jaquet-Chiffelle, David-Olivier. 'Direct and Indirect Profiling in the Light of Virtual Persons'. In *Profiling the European Citizen: Cross-Disciplinary Perspectives*, edited by Mireille Hildebrandt and Serge Gutwirth, 34–43, 2008.
- Jaume-Palasí, Lorena, and Matthias Spielkamp. 'Ethics and Algorithmic Processes for Decision Making and Decision Support', n.d., 19.

- Jefferson, Brian Jordan. 'Predictable Policing: Predictive Crime Mapping and Geographies of Policing and Race'. *Annals of the American Association of Geographers* 108, no. 1 (2 January 2018): 1–16. https://doi.org/10.1080/24694452.2017.1293500.
- Jordan, Michael I. 'Artificial Intelligence—The Revolution Hasn't Happened Yet'. *Harvard Data Science Review* 1, no. 1 (1 July 2019). https://doi.org/10.1162/99608f92.f06c6e61.
- Joseph, Matthew, Michael Kearns, Jamie Morgenstern, Seth Neel, and Aaron Roth. 'Fair Algorithms for Infinite and Contextual Bandits'. *ArXiv:1610.09559 [Cs]*, 29 October 2016. http://arxiv.org/abs/1610.09559.
- Jowell, Jeffrey. 'Is Equality a Constitutional Principle?' *Current Legal Problems* 47, no. Part_2 (1 January 1994): 1–18. https://doi.org/10.1093/clp/47.Part_2.1.
- Kamada, Yuichiro, and Fuhito Kojima. 'Fair Matching under Constraints: Theory and Applications', n.d., 55.
- Kaminski, Margot E. 'The Right to Explanation, Explained'. *Berkeley Technology Law Journal* 34 (2019): 189–218. https://doi.org/10.15779/Z38TD9N83H.
- Kammourieh, Lanah, Thomas Baar, Jos Berens, Emmanuel Letouzé, Julia Manske, John Palmer, David Sangokoya, and Patrick Vinck. 'Group Privacy in the Age of Big Data'. In *Group Privacy*, edited by Linnet Taylor, Luciano Floridi, and Bart van der Sloot, 37–66. Cham: Springer International Publishing, 2017. https://doi.org/10.1007/978-3-319-46608-8_3.
- Kant, Tanya. 'Identity, Advertising, and Algorithmic Targeting: Or How (Not) to Target Your "Ideal User"'. *MIT Case Studies in Social and Ethical Responsibilities of Computing*, no. Summer 2021 (10 August 2021). https://doi.org/10.21428/2c646de5.929a7db6.
- Kasirzadeh, Atoosa, and Andrew Smart. 'The Use and Misuse of Counterfactuals in Ethical Machine Learning'. In *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency*, 228–36. Virtual Event Canada: ACM, 2021. https://doi.org/10.1145/3442188.3445886.
- Kayser-Bril, Nicolas. 'Austria's Employment Agency Rolls out Discriminatory Algorithm, Sees No Problem'. *AlgorithmWatch* (blog), 6 October 2019. https://algorithmwatch.org/en/story/austrias-employment-agency-ams-rolls-out-discriminatory-algorithm/.
- Kelly-Lyth, Aislinn. 'Challenging Biased Hiring Algorithms'. *Oxford Journal of Legal Studies* 41, no. 4 (1 December 2021): 899–928. https://doi.org/10.1093/ojls/gqab006.
- Keyes, Os. 'The Misgendering Machines: Trans/HCI Implications of Automatic Gender Recognition'. *Proceedings of the ACM on Human-Computer Interaction* 2, no. CSCW (November 2018): 1–22. https://doi.org/10.1145/3274357.
- Khaitan, Tarunabh. *A Theory of Discrimination Law*. *A Theory of Discrimination Law*. Oxford University Press. Accessed 1 June 2020. http://www.oxfordscholarship.com/view/10.1093/acprof:oso/9780199656967.001.0001/acprof-9780199656967.
- ———. 'The Point of Discrimination Law: Securing the Freedom to Flourish'. In *The Empire of Disgust: Orejudice , Discrimination, and Policy in India and the US*, edited by Zoya Hasan, Aziz Z. Huq, Martha C. Nussbaum, and Vidhu Verma. Delhi: Oxford University Press, 2018. https://doi.org/10.1093/oso/9780199487837.003.0017.
- Kim, So Yeon. 'Les Vulnérables : Evaluating the Vulnerability Criterion in Article 14 Cases by the European Court of Human Rights'. *Legal Studies* 41, no. 4 (December 2021): 617–32. https://doi.org/10.1017/lst.2021.21.
- Kirchgaessner, Stephanie. 'Cambridge Analytica Used Data from Facebook and Politico to Help Trump'. *The Guardian*, 26 October 2017, sec. Technology. https://www.theguardian.com/technology/2017/oct/26/cambridge-analytica-used-data-from-facebook-and-politico-to-help-trump.

- Kitchin, Rob. 'Thinking Critically about and Researching Algorithms'. *Information, Communication & Society* 20, no. 1 (2017): 14–29. https://doi.org/10.1080/1369118X.2016.1154087.
- Kleinnijenhuis, Jan. 'Hoe de Belastingdienst lage inkomens profileerde in de jacht op fraude'. Trouw, 22 November 2021. https://www.trouw.nl/gs-bbb66add.
- Koops, B.-J. 'The Trouble with European Data Protection Law'. *International Data Privacy Law* 4, no. 4 (1 November 2014): 250–61. https://doi.org/10.1093/idpl/ipu023.
- Koukoulis-Spiliotopoulos, Sophia. 'The Amended Equal Treatment Directive (2002/73): An Expression of Constitutional Principles / Fundamental Rights'. *Maastricht Journal of European and Comparative Law* 12, no. 4 (December 2005): 327–68. https://doi.org/10.1177/1023263X0501200402.
- Krause, Sharon R. 'Beyond Non-Domination: Agency, Inequality and the Meaning of Freedom'. *Philosophy & Social Criticism* 39, no. 2 (1 February 2013): 187–208. https://doi.org/10.1177/0191453712470360.
- Kulk, Stefan, and Stijn van Deursen. *Juridische Aspecten van Algoritmen Die Besluiten Nemen: Een Verkennend Onderzoek.* Boom Juridisch, 2020.
- Kusner, Matt J, Joshua Loftus, Chris Russell, and Ricardo Silva. 'Counterfactual Fairness', n.d., 11.
- Kymlicka, Will, ed. *Justice in Political Philosophy: Volume 1*. Cambridge: Edward Elgar Publishing Limited, 1992.
- Laborde, Cécile. 'Political Philosophy, Social Theory, and Critical Republicanism'. In *Critical Republicanism*. Oxford: Oxford University Press, 2008. https://doi.org/10.1093/acprof:oso/9780199550210.003.0001.
- Lamont, Julian, and Christi Favor. 'Distributive Justice'. In *The Stanford Encyclopedia of Philosophy*, edited by Edward N. Zalta, Winter 2017. Metaphysics Research Lab, Stanford University, 2017. https://plato.stanford.edu/archives/win2017/entries/justice-distributive/.
- Lantschner, Emma. 'The Development of the EU Non-Discrimination Regime'. In *Reflexive Governance in EU Equality Law*. Oxford: Oxford University Press, 2021. https://doi.org/10.1093/oso/9780192843371.003.0002.
- Lavrysen, Laurens. 'Strengthening the Protection of Human Rights of Persons Living in Poverty under the ECHR'. *Netherlands Quarterly of Human Rights* 33, no. 3 (1 September 2015): 293–325. https://doi.org/10.1177/016934411503300303.
- Le Métayer, Daniel, and Julien Le Clainche. 'From the Protection of Data to the Protection of Individuals: Extending the Application of Non-Discrimination Principles'. In *European Data Protection: In Good Health?*, edited by Serge Gutwirth, Ronald Leenes, Paul De Hert, and Yves Poullet, 315–29. Dordrecht: Springer Netherlands, 2012. https://doi.org/10.1007/978-94-007-2903-2_15.
- Léchenet, Alexandre. 'A Valence-Romans, des places en crèche attribuées par un algorithme'. La Gazette des Communes, 20 August 2020. https://www.lagazettedescommunes.com/691390/a-valence-romans-des-places-en-creche-attribuees-par-un-algorithme/.
- Lee, Jocelyn R. Smith, and Michael A. Robinson. '"That's My Number One Fear in Life. It's the Police": Examining Young Black Men's Exposures to Trauma and Loss Resulting From Police Violence and Police Killings:' *Journal of Black Psychology*, 30 July 2019. https://doi.org/10.1177/0095798419865152.
- Lee, Michelle Seng Ah, Luciano Floridi, and Jatinder Singh. 'Formalising Trade-Offs beyond Algorithmic Fairness: Lessons from Ethical Philosophy and Welfare Economics'. *AI and Ethics* 1, no. 4 (November 2021): 529–44. https://doi.org/10.1007/s43681-021-00067-y.
- Lee, Min Kyung. 'Understanding Perception of Algorithmic Decisions: Fairness, Trust, and Emotion in Response to Algorithmic Management'. *Big Data & Society* 5, no. 1 (1 January 2018): 2053951718756684. https://doi.org/10.1177/2053951718756684.
- Leese, Matthias. 'The New Profiling: Algorithms, Black Boxes, and the Failure of Anti-Discriminatory Safeguards in the European Union'. *Security Dialogue* 45, no. 5 (2014): 494–511. https://doi.org/10.1177/0967010614544204.

- Legg, Andrew. *The Margin of Appreciation in International Human Rights Law: Deference and Proportionality*. Oxford Monographs in International Law. Oxford: Oxford University Press, 2012. https://doi.org/10.1093/acprof:oso/9780199650453.001.0001.
- Leijten, Ingrid. 'Defining the Scope of Economic and Social Guarantees in the Case Law of the ECtHR'. In *Shaping Rights in the ECHR: The Role of the European Court of Human Rights in Determining the Scope of Human Rights*, edited by Eva Brems and Janneke Gerards, 109–36. Cambridge: Cambridge University Press, 2014. https://doi.org/10.1017/CBO9781107337923.008.
- Lemmens, Koen. 'The Margin of Appreciation in the ECtHR's Case Law: A European Version of the Levels of Scrutiny Doctrine?' *European Journal of Law Reform* 2, no. 3 (2018): 78–96.
- Lenaerts, Koen. 'Limits on Limitations: The Essence of Fundamental Rights in the EU'. *German Law Journal* 20, no. 6 (September 2019): 779–93. https://doi.org/10.1017/glj.2019.62.
- Lerman, Jonas. 'Big Data and Its Exclusions'. *Stanford Law Review Online* 66, no. 55 (2013). https://doi.org/10.2139/ssrn.2293765.
- Lessig, Lawrence. *Code*. Version 2.0. New York: Basic Books, 2006.
- Letsas, G. 'Strasbourg's Interpretive Ethic: Lessons for the International Lawyer'. *European Journal of International Law* 21, no. 3 (1 August 2010): 509–41. https://doi.org/10.1093/ejil/chq056.
- ———. 'The Truth in Autonomous Concepts: How To Interpret the ECHR'. *European Journal of International Law* 15, no. 2 (1 April 2004): 279–305. https://doi.org/10.1093/ejil/15.2.279.
- ———. 'Two Concepts of the Margin of Appreciation'. *Oxford Journal of Legal Studies* 26, no. 4 (1 January 2006): 705–32. https://doi.org/10.1093/ojls/gql030.
- Letsas, George. *A Theory of Interpretation of the European Convention on Human Rights*. Oxford: Oxford University Press, 2007. https://doi.org/10.1093/acprof:oso/9780199203437.001.0001.
- ———. 'Human Rights, Legality, and the ECHR'. In *A Theory of Interpretation of the European Convention on Human Rights*. Oxford: Oxford University Press, 2007. https://doi.org/10.1093/acprof:oso/9780199203437.003.0002.
- Levin, Sam. 'A Beauty Contest Was Judged by AI and the Robots Didn't like Dark Skin'. *The Guardian*, 8 September 2016, sec. Technology. https://www.theguardian.com/technology/2016/sep/08/artificial-intelligence-beauty-contest-doesnt-like-black-people.
- ———. 'New AI Can Guess Whether You're Gay or Straight from a Photograph'. *The Guardian*, 7 September 2017, sec. Technology. https://www.theguardian.com/technology/2017/sep/07/new-artificial-intelligence-can-tell-whether-youre-gay-or-straight-from-a-photograph.
- Lickel, B., D. L. Hamilton, G. Wieczorkowska, A. Lewis, S. J. Sherman, and A. N. Uhles. 'Varieties of Groups and the Perception of Group Entitativity.' *Journal of Personality and Social Psychology* 78, no. 2 (2000): 223–46.
- Lindblom, Lars. 'In Defense of Rawlsian Fair Equality of Opportunity'. *Philosophical Papers* 47, no. 2 (4 May 2018): 235–63. https://doi.org/10.1080/05568641.2018.1445550.
- Lipper-Rasmussen, Kasper, ed. *The Routledge Handbook of the Ethics of Discrimination*. London and New York: Routledge, 2018.
- Lippert-Rasmussen, Kasper. *Born Free and Equal?: A Philosophical Inquiry Into the Nature of Discrimination*. OUP USA, 2014.
- ———. *Discrimination*. Edited by Serena Olsaretti. Vol. 1. Oxford University Press, 2018. https://doi.org/10.1093/oxfordhb/9780199645121.013.20.
- ———. 'Discrimination: An Intriguing but Underexplored Issue in Ethics and Political Philosophy'. *Moral Philosophy and Politics; Berlin* 2, no. 2 (2015): 207–17. http://dx.doi.org.kuleuven.ezproxy.kuleuven.be/10.1515/mopp-2015-0026.

- Lippert-rasmussen, Kasper. 'The Badness of Discrimination'. *Ethical Theory and Moral Practice* 9, no. 2 (April 2006): 167–85. https://doi.org/10.1007/s10677-006-9014-x.
- Lippert-Rasmussen, Kasper, ed. *The Routledge Handbook of the Ethics of Discrimination*. 1st ed. Routledge Handbooks. New York: Routledge, 2018. https:///bookshop/product/The-Routledge-Handbook-of-the-Ethics-of-Discrimination-by-Kasper-Lippert-Rasmussen-editor/9781317400769.
- Lippmann, Walter. *Public Opinion*. New York, Harcourt, Brace and Company, 1922. http://archive.org/details/publicopinion00lippgoog.
- Liu, Han, Alexander Gegov, and Mihaela Cocea. *Rule Based Systems for Big Data*. Vol. 13. Studies in Big Data. Cham: Springer International Publishing, 2016. https://doi.org/10.1007/978-3-319-23696-4.
- Liu, Lydia T., Sarah Dean, Esther Rolf, Max Simchowitz, and Moritz Hardt. 'Delayed Impact of Fair Machine Learning'. *ArXiv:1803.04383 [Cs, Stat]*, 12 March 2018. http://arxiv.org/abs/1803.04383.
- Lohr, Steve. 'Banking Start-Ups Adopt New Tools for Lending'. *The New York Times*, 19 January 2015, sec. Technology. https://www.nytimes.com/2015/01/19/technology/banking-start-ups-adopt-new-tools-for-lending.html.
- Lopez, Paola. 'Bias Does Not Equal Bias: A Socio-Technical Typology of Bias in Data-Based Algorithmic Systems'. *Internet Policy Review* 10, no. 4 (7 December 2021). https://policyreview.info/articles/analysis/bias-does-not-equal-bias-socio-technical-typology-bias-data-based-algorithmic.
- ———. 'Reinforcing Intersectional Inequality via the AMS Algorithm in Austria'. In *Proceedings of the STS Conference Graz*, 21. Graz, 2019.
- Lulamae, Josephine. 'Automated Society: Are Green Influencers Being "Shadowbanned" on Instagram?' Accessed 11 February 2022. https://r.algorithmwatch.org/nl2/x3ysn/158g.html?m=AL8AAL79d5EAAAAPjTIAAAZJGn8AAAAA6uQAAB1gABB0KQBiBfudOgjIB3G0TWetucmFgzeQTAAQJ0I&b=bfbfe00e&e=3d32336c&x=0A1rc6D5oqSxnwruhXsAaIRdCFsr9ZW1dtSNvoBxFTE.
- Lundgard, Alan. 'Measuring Justice in Machine Learning'. *Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency*, 27 January 2020, 680–680. https://doi.org/10.1145/3351095.3372838.
- Lurie, Guy. 'Proportionality and the Right to Equality'. *German Law Journal* 21, no. 2 (February 2020): 174–96. https://doi.org/10.1017/glj.2020.8.
- Macleod, C. W. 'Form and Meaning in the Melian Dialogue'. *Historia: Zeitschrift Für Alte Geschichte* 23, no. 4 (1974): 385–400.
- Mahdawi, Arwa. 'This AI-Powered App Will Tell You If You're Beautiful – and Reinforce Biases, Too'. *The Guardian*, 6 March 2021, sec. Opinion. https://www.theguardian.com/commentisfree/2021/mar/06/ai-powered-app-tell-you-beautiful-reinforce-biases.
- Mahlmann, Matthias. 'ECJ Headscarf Series (3): The Everyday Troubles of Pluralism'. *Strasbourg Observers* (blog), 12 September 2016. https://strasbourg.weichie.dev/2016/09/12/ecj-headscarf-series-3-the-everyday-troubles-of-pluralism/.
- Makkonen, Timo. 'Multiple, Compound and Intersectional Discrimination: Bringing the Experiences of the Most Marginalized to the Fore.' Turku: Åbo Akademi University. Institute for Human Rights, 2002.
- Malgieri, Gianclaudio. 'Automated Decision-Making in the EU Member States: The Right to Explanation and Other "Suitable Safeguards" in the National Legislations'. *Computer Law & Security Review* 35, no. 5 (1 October 2019): 105327. https://doi.org/10.1016/j.clsr.2019.05.002.
- Malgieri, Gianclaudio, and Giovanni Comandé. 'Why a Right to Legibility of Automated Decision-Making Exists in the General Data Protection Regulation'. *International Data Privacy Law* 7, no. 4 (1 November 2017): 243–65. https://doi.org/10.1093/idpl/ipx019.

- Malleson, Kate. 'Equality Law and the Protected Characteristics'. *The Modern Law Review* 81, no. 4 (2018): 598–621. https://doi.org/10.1111/1468-2230.12353.
- ———. 'Equality Law and the Protected Characteristics: Equality Law and the Protected Characteristics'. *The Modern Law Review* 81, no. 4 (July 2018): 598–621. https://doi.org/10.1111/1468-2230.12353.
- Manders-Huits, Noëmi. 'Practical Versus Moral Identities in Identity Management'. *Ethics and Information Technology* 12, no. 1 (2010): 43–55. https://doi.org/10.1007/s10676-010-9216-8.
- ———. 'Practical versus Moral Identities in Identity Management'. *Ethics and Information Technology* 12, no. 1 (March 2010): 43–55. https://doi.org/10.1007/s10676-010-9216-8.
- Mann, Gideon, and Cathy O'Neil. 'Hiring Algorithms Are Not Neutral'. *Harvard Business Review*, 9 December 2016. https://hbr.org/2016/12/hiring-algorithms-are-not-neutral.
- Mann, Monique, and Tobias Matzner. 'Challenging Algorithmic Profiling: The Limits of Data Protection and Anti-Discrimination in Responding to Emergent Discrimination'. *Big Data & Society* 6, no. 2 (1 July 2019): 2053951719895805. https://doi.org/10.1177/2053951719895805.
- Mantelero, Alessandro. 'From Group Privacy to Collective Privacy: Towards a New Dimension of Privacy and Data Protection in the Big Data Era'. In *Group Privacy: New Challenges of Data Technologies*, edited by Linnet Taylor, Luciano Floridi, and Bart van der Sloot, 139–58. Cham: Springer International Publishing, 2017. https://doi.org/10.1007/978-3-319-46608-8_8.
- Martin, Denis. 'Article 20 CFR'. In *Article 20 CFR*. Oxford University Press, 2019. https://doi.org/10.1093/oso/9780198759393.003.541.
- Mason, Andrew. 'Rawlsian Fair Equality of Opportunity'. In *Levelling the Playing Field*. Oxford: Oxford University Press, 2006. https://doi.org/10.1093/acprof:oso/9780199264414.003.0004.
- Mattu, Surya, Jeff Larson, Lauren Kirchner, and Julia Angwin. 'How We Analyzed the COMPAS Recidivism Algorithm'. ProPublica, 23 May 2016. https://www.propublica.org/article/how-we-analyzed-the-compas-recidivism-algorithm?token=e3k6mE9frXgAOt3zvKfhyLDrUGgrhmOb.
- Matzat, Lorenz, and Minna Ruckenstein. 'Finnish Credit Score Ruling Raises Questions about Discrimination and How to Avoid It'. *AlgorithmWatch* (blog), 21 November 2018. https://algorithmwatch.org/en/finnish-credit-score-ruling-raises-questions-about-discrimination-and-how-to-avoid-it/.
- McBride, Jeremy. 'Proportionality and the European Convention on Human Rights.' In *The Principle of Proportionality in the Laws of Europe*, edited by Evelyn Ellis, 23–36. Hart Publishing, 1999. https://doi.org/10.5040/9781472561947.
- McColgan, Aileen. 'Defining the Protected Characteristics'. In *Discrimination, Equality and the Law*, 1st ed., 38–69. London: Hart Publishing, 2014. http://www.bloomsburycollections.com/book///.
- McCrudden, C. 'Human Dignity and Judicial Interpretation of Human Rights'. *European Journal of International Law* 19, no. 4 (1 September 2008): 655–724. https://doi.org/10.1093/ejil/chn043.
- McCrudden, Christopher. *Buying Social Justice: Equality, Government Procurement, and Legal Change*. Oxford University Press, 2007. https://doi.org/10.1093/acprof:oso/9780199232420.001.0001.
- McCrudden, Christopher, and Haris Kountouros. 'Human Rights and European Equality Law'. In *Equality Law in an Enlarged European Union: Understanding the Article 13 Directives*, edited by Helen Meenan, 73–116. Cambridge: Cambridge University Press, 2007. https://doi.org/10.1017/CBO9780511493898.004.
- McCrudden, Christopher, and Sacha Prechal. 'The Concepts of Equality and Non-Discrimination in Europe: A Practical Approach'. European Network of Legal Experts in the Field of Gender Equality, November 2009.
- Medina, Lucas Waldron, Brenda. 'TSA's Body Scanners Are Gender Binary. Humans Are Not.' ProPublica. Accessed 7 February 2022. https://www.propublica.org/article/tsa-transgender-travelers-scanners-invasive-searches-often-wait-on-the-other-side?token=7bjY-MRzWk5Ed4DCZRvFVYwt2HBrAFXd.

- Mehrabi, Ninareh, Fred Morstatter, Nripsuta Saxena, Kristina Lerman, and Aram Galstyan. 'A Survey on Bias and Fairness in Machine Learning'. *ArXiv:1908.09635 [Cs]*, 17 September 2019. http://arxiv.org/abs/1908.09635.
- Metz, Cade. 'In Major AI Breakthrough, Google System Secretly Beats Top Player at the Ancient Game of Go'. *Wired*. Accessed 19 May 2022. https://www.wired.com/2016/01/in-a-huge-breakthrough-googles-ai-beats-a-top-player-at-the-game-of-go/.
- Miceli, Milagros. 'AI's Symbolic Power: Classification in the Age of Automation', 2019.
- Miceli, Milagros, and Julian Posada. 'A Question of Power: How Task Instructions Shape Training Data'. Tagungsbeitrag presented at the Symposium on Biases in Human Computation and Crowdsourcing. BHCC2020, 11 November 2020. https://sites.google.com/sheffield.ac.uk/bhcc-2020/program?authuser=0.
- Miceli, Milagros, Julian Posada, and Tianling Yang. 'Studying Up Machine Learning Data: Why Talk About Bias When We Mean Power?' *ArXiv:2109.08131 [Cs]*, 16 September 2021. http://arxiv.org/abs/2109.08131.
- Miceli, Milagros, Martin Schuessler, and Tianling Yang. 'Between Subjectivity and Imposition: Power Dynamics in Data Annotation for Computer Vision'. *Proceedings of the ACM on Human-Computer Interaction* 4, no. CSCW2 (14 October 2020): 1–25. https://doi.org/10.1145/3415186.
- Miceli, Milagros, Tianling Yang, Laurens Naudts, Martin Schuessler, Diana Serbanescu, and Alex Hanna. 'Documenting Computer Vision Datasets: An Invitation to Reflexive Data Practices'. In *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency*, 161–72. Virtual Event Canada: ACM, 2021. https://doi.org/10.1145/3442188.3445880.
- Miller, David. 'Justice'. In *The Stanford Encyclopedia of Philosophy*, edited by Edward N. Zalta, Fall 2017. Metaphysics Research Lab, Stanford University, 2017. https://plato.stanford.edu/archives/fall2017/entries/justice/.
- Miller, Tim. 'Explanation in Artificial Intelligence: Insights from the Social Sciences'. *ArXiv:1706.07269 [Cs]*, 22 June 2017. http://arxiv.org/abs/1706.07269.
- Milmo, Dan, and Dan Milmo Global technology editor. 'Twitter Admits Bias in Algorithm for Rightwing Politicians and News Outlets'. *The Guardian*, 22 October 2021, sec. Technology. https://www.theguardian.com/technology/2021/oct/22/twitter-admits-bias-in-algorithm-for-rightwing-politicians-and-news-outlets.
- Mittelstadt, Brent. 'From Individual to Group Privacy in Big Data Analytics'. *Philosophy & Technology* 30, no. 4 (December 2017): 475–94. https://doi.org/10.1007/s13347-017-0253-7.
- Mittelstadt, Brent Daniel, Patrick Allo, Mariarosaria Taddeo, Sandra Wachter, and Luciano Floridi. 'The Ethics of Algorithms: Mapping the Debate'. *Big Data & Society* 3, no. 2 (December 2016): 1–21. https://doi.org/10.1177/2053951716679679.
- Mittelstadt, Brent, Chris Russell, and Sandra Wachter. 'Explaining Explanations in AI'. In *Proceedings of the Conference on Fairness, Accountability, and Transparency*, 279–88. FAT* '19. New York, NY, USA: Association for Computing Machinery, 2019. https://doi.org/10.1145/3287560.3287574.
- Moller, K. 'Proportionality: Challenging the Critics'. *International Journal of Constitutional Law* 10, no. 3 (1 July 2012): 709–31. https://doi.org/10.1093/icon/mos024.
- Molnar, Christoph. *Interpretable Machine Learning*. Accessed 8 September 2022. https://christophm.github.io/interpretable-ml-book/.
- Morgenroth, Thekla, and Michelle K. Ryan. 'The Effects of Gender Trouble: An Integrative Theoretical Framework of the Perpetuation and Disruption of the Gender/Sex Binary'. *Perspectives on Psychological Science* 16, no. 6 (1 November 2021): 1113–42. https://doi.org/10.1177/1745691620902442.
- Muir, Elise. *EU Equality Law: The First Fundamental Rights Policy of the EU*. Oxford Studies in European Law. Oxford: Oxford University Press, 2018. https://doi.org/10.1093/oso/9780198814665.001.0001.

- ———. *EU Equality Law: The First Fundamental Rights Policy of the EU*. Oxford Studies in European Law. Oxford: Oxford University Press, 2018. https://doi.org/10.1093/oso/9780198814665.001.0001.
- ———. 'The Essence of the Fundamental Right to Equal Treatment: Back to the Origins'. *German Law Journal* 20, no. 6 (September 2019): 817–39. https://doi.org/10.1017/glj.2019.64.
- ———. 'The Transformative Function of EU Equality Law'. *European Review of Private Law* 21, no. 5/6 (1 November 2013). https://kluwerlawonline.com/journalarticle/European+Review+of+Private+Law/21.5/ERPL2013075.
- Muldoon, James, and Paul Raekstad. 'Algorithmic Domination in the Gig Economy'. *European Journal of Political Theory*, 23 February 2022, 14748851221082078. https://doi.org/10.1177/14748851221082078.
- Murphy, Colleen, and Paolo Gardoni. 'Design, Risk and Capabilities'. In *The Capability Approach, Technology and Design*, edited by Ilse Oosterlaken and Jeroen van den Hoven, 173–88. Philosophy of Engineering and Technology. Dordrecht: Springer Netherlands, 2012. https://doi.org/10.1007/978-94-007-3879-9_10.
- Murphy, Heather. 'Why Stanford Researchers Tried to Create a "Gaydar" Machine'. *The New York Times*, 9 October 2017, sec. Science. https://www.nytimes.com/2017/10/09/science/stanford-sexual-orientation-study.html.
- Nagel, Thomas. 'John Rawls and Affirmative Action'. *The Journal of Blacks in Higher Education*, no. 39 (2003): 82. https://doi.org/10.2307/3134387.
- Nathan, Christopher. 'What Is Basic Equality?' In *Do All Persons Have Equal Moral Worth?*, edited by Uwe Steinhoff, 1–16. Oxford University Press, 2014. https://doi.org/10.1093/acprof:oso/9780198719502.003.0001.
- National Center for Transgender Equality. 'Airport Security'. *National Center for Transgender Equality*. Accessed 29 April 2022. https://transequality.org/know-your-rights/airport-security.
- Naudts, Laurens. 'AI En Algoritmische Categorisatie: Gelijkheid En Non-Discriminatie'. In *Artificiële Intelligentie En Maatschappij*, edited by Jan De Bruyne and Nicolas Bouteca, 223–47. Oud-Turnhout/'s-Hertogenbosch: Gompel & Svacina, 2021.
- ———. 'Criminal Profiling and Non-Discrimination: On Firm Grounds for the Digital Era?' In *Security and Law. Legal and Ethical Aspects of Public Security, Cyber Security and Critical Infrastructure Security.*, edited by Anton Vedder, Jessica Schroers, Charlotte Ducuing, and Peggy Valcke, 63–96. Camdridge, Antwerp, Chicago: Intersentia, 2019.
- ———. 'Fair or Unfair Differentiation? Luck Egalitarianism as a Lens for Evaluating Algorithmic Decision-Making.' London, 2017.
- ———. 'How Machine Learning Generates Unfair Inequalities and How Data Protection Instruments May Help in Mitigating Them'. In *Data Protection and Privacy: The Internet of Bodies*, edited by Ronald Leenes, Serge Gutwirth, Paul De Hert, and Rosamunde van Brakel, 71–92. Hart Publishing, 2018.
- ———. 'Towards Accountability: The Articulation and Formalization of Fairness in Machine Learning'. *SSRN Electronic Journal*, 2018. https://doi.org/10.2139/ssrn.3298847.
- Naudts, Laurens, Pierre Dewitte, and Jef Ausloos. 'Meaningful Transparency through Data Rights: A Multidimensional Analysis'. In *Research Handbook on EU Data Protection Law*, 530–71. Edward Elgar Publishing Limited, 2022.
- Neuvonen, Päivi Johanna. '"Inequality in Equality" in the European Union Equality Directives: A Friend or a Foe of More Systematized Relationships between the Protected Grounds?' *International Journal of Discrimination and the Law* 15, no. 4 (December 2015): 222–40. https://doi.org/10.1177/1358229115591412.
- ILOSTAT. 'New ILO Database Highlights Labour Market Challenges of Persons with Disabilities', 13 June 2022. https://ilostat.ilo.org/new-ilo-database-highlights-labour-market-challenges-of-persons-with-disabilities/.

- Nicas, Jack, Raymond Zhong, and Daisuke Wakabayashi. 'Censorship, Surveillance and Profits: A Hard Bargain for Apple in China'. *The New York Times*, 17 May 2021, sec. Technology. https://www.nytimes.com/2021/05/17/technology/apple-china-censorship-data.html.
- Nikolaidis, Charilaos. 'Rethinking Likeness and Comparability in Equality Claims Brought before the European Court of Human Rights'. *Public Law*, July 2020, 448–67.
- Nilsson, Anna. 'Same, Same but Different: Proportionality Assessments and Equality Norms'. *Oslo Law Review* 7, no. 3 (17 December 2020): 126–44. https://doi.org/10.18261/ISSN.2387-3299-2020-03-01.
- Nissenbaum, Helen. 'Computing and Accountability', n.d., 9.
- Noble, Safiya Umoja. *Algorithms of Oppression: How Search Engines Reinforce Racism*. *Algorithms of Oppression*. New York University Press, 2018. https://doi.org/10.18574/9781479833641.
- Northpointe. 'COMPAS Risk & Need Assessment System'. Northpointe, 2012. http://www.northpointeinc.com/files/downloads/FAQ_Document.pdf.
- Nossek, Alexa, and Julia Belz. 'Notes on Pettit's Concept of Orthonomy'. In *Philip Pettit: Five Themes from His Work*, edited by Simon Derpmann and David P. Schweikard, 61–69. Münster Lectures in Philosophy. Cham: Springer International Publishing, 2016. https://doi.org/10.1007/978-3-319-26103-4_6.
- Ntoutsi, Eirini, Pavlos Fafalios, Ujwal Gadiraju, Vasileios Iosifidis, Wolfgang Nejdl, Maria-Esther Vidal, Salvatore Ruggieri, et al. "Bias in Data-Driven Artificial Intelligence Systems—An Introductory Survey." WIREs Data Mining and Knowledge Discovery 10, no. 3 (2020): e1356. https://doi.org/10.1002/widm.1356.
- Nussbaum, Martha. 'Capabilities as Fundamental Entitlements: Sen and Social Justice'. *Feminist Economics* 9, no. 2–3 (January 2003): 33–59. https://doi.org/10.1080/1354570022000077926.
- ———. 'Human Dignity and Political Entitlements'. *Human Dignity and Bioethics: Essays Commissioned by the President's Council on Bioethics*, 1 January 2008. https://chicagounbound.uchicago.edu/book_chapters/202.
- ———. 'Nature, Function, and Capability: Aristotle on Political Distribution'. *Oxford Studies in Ancient Philosophy: Supplementary Volume*, 1 January 1988. https://chicagounbound.uchicago.edu/book_chapters/782.
- ———. 'The Capabilities of People with Cognitive Disabilities'. *Metaphilosophy* 40, no. 3/4 (2009): 331–51.
- ———. 'Women and Equality: The Capabilities Approach Special Issue: Women, Gender and Work (Part I)'. *International Labour Review* 138, no. 3 (1999): 227–46.
- Nussbaum, Martha C. 'Ageing, Stigma, and Disgust'. In *The Empire of Disgust*. Delhi: Oxford University Press, 2018. https://doi.org/10.1093/oso/9780199487837.003.0008.
- ———. 'Aristotle, Politics, and Human Capabilities: A Response to Antony, Arneson, Charlesworth, and Mulgan'. *Ethics* 111, no. 1 (2000): 102–40. https://doi.org/10.1086/233421.
- ———, ed. 'Compassion and Public Life'. In *Upheavals of Thought: The Intelligence of Emotions*, 401–54. Cambridge: Cambridge University Press, 2001. https://doi.org/10.1017/CBO9780511840715.010.
- ———, ed. 'Compassion: The Philosophical Debate'. In *Upheavals of Thought: The Intelligence of Emotions*, 354–400. Cambridge: Cambridge University Press, 2001. https://doi.org/10.1017/CBO9780511840715.009.
- ———. *Creating Capabilities: The Human Development Approach*. Cambridge, Massachusetts ; London, England: The Belknap Press of Harvard University Press, 2011.
- ———. *Frontiers of Justice: Disability, Nationality, Species Membership*. The Tanner Lectures on Human Values. Cambridge, Massachusetts, London: Belknap Press: Harvard University Press, 2006.
- ———. *Frontiers of Justice: Disability, Nationality, Species Membership*. Cambridge, Massachusetts ; London, England: Belknap Press: Harvard University Press, 2006.
- ———. 'Human Functioning and Social Justice: In Defense of Aristotelian Essentialism'. *Political Theory* 20, no. 2 (1992): 202–46.

- ———. 'The Capabilities Approach and the History of Philosophy'. In *The Cambridge Handbook of the Capability Approach*, edited by Enrica Chiappero-Martinetti, Siddiqur Osmani, and Mozaffar Qizilbash, 1st ed., 13–39. Cambridge University Press, 2020. https://doi.org/10.1017/9781316335741.003.
- ———. *The Cosmopolitan Tradition: A Noble But Flawed Ideal*. Cambridge, Massachusetts ; London, England: Belknap Press: Harvard University Press, 2019.
- ———. 'The Supreme Court 2006 Term. Foreword: Constitutions and Capabilities: "Perceptions" Against Lofty Formalism.' *Harvard Law Review* 121 (2006): 94.
- ———. *Women and Human Development: The Capabilities Approach*. 12th ed. New York: Cambridge University Press, 2000.
- Obermeyer, Ziad, Brian Powers, Christine Vogeli, and Sendhil Mullainathan. 'Dissecting Racial Bias in an Algorithm Used to Manage the Health of Populations'. *Science* 366, no. 6464 (25 October 2019): 447–53. https://doi.org/10.1126/science.aax2342.
- O'Connell, Rory. 'Cinderella Comes to the Ball: Art 14 and the Right to Non-Discrimination in the ECHR'. *Legal Studies* 29 (2009): 211–29.
- ———, ed. 'European Principles of Interpretation and Application'. In *Law, Democracy and the European Court of Human Rights*, 65–83. Cambridge: Cambridge University Press, 2020. https://doi.org/10.1017/9781139547246.006.
- Ohlström, Marcus, Marco Solinas, and Olivier Voirol. 'Nancy Fraser and Axel Honneth's Redistribution or Recognition? A Political-Philosophical Exchange1'. *Iris* 3, no. 5 (2011): 205–21.
- O'Neil, Cathy. *Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy*. New York, NY, USA: Crown Publishing Group, 2016.
- O'Neill, Martin. 'What Should Egalitarians Believe?' *Philosophy & Public Affairs* 36, no. 2 (March 2008): 119–56. https://doi.org/10.1111/j.1088-4963.2008.00130.x.
- Oosterlaken, Ilse. 'Design for Development: A Capability Approach'. *Design Issues* 25, no. 4 (October 2009): 91–102. https://doi.org/10.1162/desi.2009.25.4.91.
- O'Shea, Tom. 'Disability and Domination: Lessons from Republican Political Philosophy'. *Journal of Applied Philosophy* 35, no. 1 (2018): 133–48. https://doi.org/10.1111/japp.12149.
- Ostrovsky, Aaron A. 'What's So Funny about Peace, Love, and Understanding - How the Margin of Appreciation Doctrine Preserves Core Human Rights within Cultural Diversity and Legitimises International Human Rights Tribunals International Law'. *Hanse Law Review* 1, no. 1 (2005): 47–64.
- Ouald-Chaib, Saïla, and Valeska David. 'European Court of Justice Keeps the Door to Religious Discrimination in the Private Workplace Opened. The European Court of Human Rights Could Close It.' *Strasbourg Observers* (blog), 2017. https://strasbourg.weichie.dev/2017/03/27/european-court-of-justice-keeps-the-door-to-religious-discrimination-in-the-private-workplace-opened-the-european-court-of-human-rights-could-close-it/.
- Pasquale, Frank. *The Black Box Society: The Secret Algorithms That Control Money and Information*. First Harvard University Press paperback edition. Cambridge, Massachusetts London, England: Harvard University Press, 2015.
- Perez, Caroline Criado. *Invisible Women: Exposing Data Bias in a World Designed for Men*. 01 ed. London: Chatto & Windus, 2019.
- ———. 'The Deadly Truth about a World Built for Men – from Stab Vests to Car Crashes'. *The Guardian*, 23 February 2019, sec. Life and style. https://www.theguardian.com/lifeandstyle/2019/feb/23/truth-world-built-for-men-car-crashes.
- Peroni, L., and A. Timmer. 'Vulnerable Groups: The Promise of an Emerging Concept in European Human Rights Convention Law'. *International Journal of Constitutional Law* 11, no. 4 (1 October 2013): 1056–85. https://doi.org/10.1093/icon/mot042.

- Peroni, Lourdes. 'Age and Gender Discrimination: Laudable Anti-Stereotyping Reasoning in Carvalho Pinto V. Portugal'. *Strasbourg Observers* (blog), 28 September 2017. https://strasbourgobservers.com/2017/09/28/age-and-gender-discrimination-laudable-anti-stereotyping-reasoning-in-carvalho-pinto-v-portugal/.
- Pettit, Philip. *Judging Justice. An Introduction to Contemporary Political Philosophy*. London, Boston and Henley: Routlegde & Kegan Paul, 1980.
- ———. *On The People's Terms: A Republican Theory and Model of Democracy*. Cambridge, New York, Melbourne, Madrid, Cape Town, Singapore, Sao Paolo, Delhi, Mexico City: Cambridge University Press, 2012.
- ———. *Republicanism: A Theory of Freedom and Government*. *Republicanism*. Oxford University Press, 1999. https://www.oxfordscholarship.com/view/10.1093/0198296428.001.0001/acprof-9780198296423.
- Pettit, Philip, and Michael Smith. 'Backgrounding Desire'. *The Philosophical Review* 99, no. 4 (October 1990): 565. https://doi.org/10.2307/2185617.
- Pierik, R. H. M. 'The Ideal of Equality in Political Philosophy'. In *The Importance of Ideals*, edited by Wibren van der Burg and Sanne Taekema, 173–96. Brussels, Oxford, etc.: Peter Lang, 2004. https://repository.ubn.ru.nl/handle/2066/67454.
- Pogge, Thomas. *John Rawls: His Life and Theory of Justice*. Translated by Michelle Kosch. New York: Oxford University Press, 2007.
- ———. 'Relational Conceptions of Justice: Responsibilities for Health Outcomes'. In *Public Health, Ethics, and Equity*, edited by Sudhir Anand, Fabienne Peter, and Amartya Sen, 135–61. Oxford University Press, 2004.
- Pogge, Thomas W. 'Relational Conceptions of Justice: Responsibilities for Health Outcomes'. *Veritas* 46, no. 1 (1 January 2001): 51–75.
- ———. 'Three Problems with Contractarian-Consequentialist Ways of Assessing Social Institutions*: THOMAS W. POGGE'. *Social Philosophy and Policy* 12, no. 2 (1995): 241–66. https://doi.org/10.1017/S026505250000474X.
- Prechal, Sacha. 'Equality of Treatment, Non-Discrimination and Social Policy: Achievements in Three Times'. *Common Market Law Review*, no. 41 (2004): 533–51.
- ———. 'Horizontal Direct Effect of the Charter of Fundamental Rights of the EU'. *Revista de Derecho Comunitario Europeo*, no. 66 (n.d.): 407–26. https://doi.org/10.18042/cepc/rdce.66.04.
- ———. 'Non-Discrimination Does Not Fall Down From Heaven: The Context and Evolution of Non-Discrimination in EU Law'. *Eric Stein Working Paper*, Eric Stein Working Paper, 2009, no. 4 (2009): 17.
- 'Principles of Biomedical Ethics - Paperback - Tom L. Beauchamp; James F. Childress - Oxford University Press'. Accessed 13 July 2020. https://global.oup.com/ushe/product/principles-of-biomedical-ethics-9780190640873?cc=be&lang=en&.
- 'Profiling the Poor in the Dutch Welfare State - NYU School of Law – CHRGJ'. Accessed 3 August 2022. https://chrgj.org/2019/11/01/profiling-the-poor-in-the-dutch-welfare-state/.
- Rachovitsa, Adamantia, and Niclas Johann. 'The Human Rights Implications of the Use of AI in the Digital Welfare State: Lessons Learned from the Dutch SyRI Case'. *Human Rights Law Review* 22, no. 2 (1 June 2022): ngac010. https://doi.org/10.1093/hrlr/ngac010.
- Ramirez, Edith. 'The Privacy Challenges of Big Data: A View from the Lifeguard's Chair'. In *Keynote Address*. Aspen, 2013. https://www.ftc.gov/sites/default/files/documents/public_statements/privacy-challenges-big-data-view-lifeguard%E2%80%99s-chair/130819bigdataaspen.pdf.
- Ramirez, Edith, Julie Brill, Maureen K. Ohlhausen, Joshua D. Wright, and Terrell McSweeny. 'Data Brokers: A Call for Transparency and Accountability'. Federal Trade Commission, 2014.
- Rawls, John. 'A Kantian Conception of Equality'. *Cambridge Review*, 1975, 94–99.
- ———. *A Theory of Justice*. Cambridge, Massachusetts, London: Harvard University Press (Belknap Press), 1971.

- ———. *A Theory of Justice: Revised Edition*. Harvard University Press, 1999. https://doi.org/10.2307/j.ctvkjb25m.
- ———. 'A Well-Ordered Society'. In *Philosophy, Politics and Society*, edited by Peter Laslett and James S. Fishkin, 5:6–20. Oxford, UK: Blackwell Publishers, 1979.
- ———. 'Fairness to Goodness'. *The Philosophical Review* 84, no. 4 (October 1975): 536. https://doi.org/10.2307/2183853.
- ———. *Justice as Fairness: A Restatement*. Edited by Kelly Erin. Cambridge, Massachusetts ; London, England: Belknap Press: Harvard University Press, 2001.
- ———. *Political Liberalism*. New York, Chichester, West Sussex: Columbia University Press, 1993.
- ———. 'The Basic Liberties and Their Priority'. edited by Sterling M Mcmurrin, 89. The Tanner Lectures on Human Values, 1981.
- Réaume, Denise. 'Dignity, Equality, and Comparison'. In *Philosophical Foundations of Discrimination Law*. Oxford: Oxford University Press, 2013. https://doi.org/10.1093/acprof:oso/9780199664313.003.0002.
- Réaume, Denise G. 'Discrimination and Dignity'. *Louisiana Law Review* 63, no. 3 (2003): 53.
- Richardson, Rashida, Jason M. Schultz, and Kate Crawford. 'Dirty Data, Bad Predictions: How Civil Rights Violations Impact Police Data, Predictive Policing Systems, and Justice'. *New York University Law Review Online* 94 (2019): 15–55.
- Richter, Dagmar. 'Overweight and Obesity as Novel Grounds of Discrimination'. In *The European Union as Protector and Promoter of Equality*, edited by Thomas Giegerich, 297–335. European Union and Its Neighbours in a Globalized World. Cham: Springer International Publishing, 2020. https://doi.org/10.1007/978-3-030-43764-0_16.
- Riekeles, Georg. 'I Saw First-Hand How US Tech Giants Seduced the EU – and Undermined Democracy'. *The Guardian*, 28 June 2022, sec. Opinion. https://www.theguardian.com/commentisfree/2022/jun/28/i-saw-first-hand-tech-giants-seduced-eu-google-meta.
- Roberge, Jonathan, and Louis Melançon. 'Being the King Kong of Algorithmic Culture Is a Tough Job after All: Google's Regimes of Justification and the Meanings of Glass'. *Convergence* 23, no. 3 (1 June 2017): 306–24. https://doi.org/10.1177/1354856515592506.
- Robeyns, Ingrid. 'Capabilitarianism'. *Journal of Human Development and Capabilities* 17, no. 3 (2 July 2016): 397–414. https://doi.org/10.1080/19452829.2016.1145631.
- ———. 'The Capability Approach'. In *The Stanford Encyclopedia of Philosophy*, edited by Edward N. Zalta, Winter 2016. Metaphysics Research Lab, Stanford University, 2016. https://plato.stanford.edu/archives/win2016/entries/capability-approach/.
- ———. 'The Capability Approach in Practice*'. *Journal of Political Philosophy* 14, no. 3 (2006): 351–76. https://doi.org/10.1111/j.1467-9760.2006.00263.x.
- ———. *Wellbeing, Freedom and Social Justice: The Capability Approach Re-Examined*. Open Book Publishers, 2017. https://doi.org/10.11647/OBP.0130.
- Robeyns, Ingrid, and Harry Brighouse. 'Introduction: Social Primary Goods and Capabilities as Metrics of Justice'. In *Measuring Justice*, edited by Harry Brighouse and Ingrid Robeyns, 1–14. Cambridge: Cambridge University Press, 2010. https://doi.org/10.1017/CBO9780511810916.001.
- Rouvroy, Antoinette, and Yves Poullet. 'The Right to Informational Self-Determination and the Value of Self-Development: Reassessing the Importance of Privacy for Democracy'. In *Reinventing Data Protection*, 45–76. Dordrecht: Springer, 2009.

- Sánchez-Monedero, Javier, and Lina Dencik. 'The Politics of Deceptive Borders: "Biomarkers of Deceit" and the Case of IBorderCtrl'. *Information, Communication & Society* 25, no. 3 (17 February 2022): 413–30. https://doi.org/10.1080/1369118X.2020.1792530.
- Sanders, Elizabeth B.-N., and Pieter Jan Stappers. 'Co-Creation and the New Landscapes of Design'. *CoDesign* 4, no. 1 (1 March 2008): 5–18. https://doi.org/10.1080/15710880701875068.
- Sangiovanni, Andrea. *Humanity without Dignity: Moral Equality, Respect, and Human Rights.* Harvard University Press, 2017.
- Schauer, Frederick. *Profiles, Probabilities and Stereotypes*. Cambridge, Massachusetts; London: The Belknap Press of Harvard University Press, 2003.
- Scheffler, Samuel. 'The Practice of Equality'. In *Social Equality: On What It Means to Be Equals*, edited by Carina Fourie, Fabian Schuppert, and Ivo Wallimann-Helmer, 21–44. New York: Oxford University Press, 2015.
- ———. 'What Is Egalitarianism?' *Philosophy & Public Affairs* 31, no. 1 (2003): 5–39. https://doi.org/10.1111/j.1088-4963.2003.00005.x.
- ———. 'What Is Egalitarianism?' *Philosophy & Public Affairs* 31, no. 1 (2003): 5–39.
- Schemmel, Christian. 'Distributive and Relational Equality': *Politics, Philosophy & Economics*, 15 September 2011. https://doi.org/10.1177/1470594X11416774.
- Schepman, Thibaut. 'Gros pistons et petites manip pour une place en crèche à Boulogne-Billancourt'. BuzzFeed. Accessed 3 October 2022. https://www.buzzfeed.com/fr/thibautschepman/gros-pistons-et-petites-manip-pour-une-place-en-creche-a-bou.
- Scheuerman, Morgan Klaus, Jacob M. Paul, and Jed R. Brubaker. 'How Computers See Gender: An Evaluation of Gender Classification in Commercial Facial Analysis Services'. *Proceedings of the ACM on Human-Computer Interaction* 3, no. CSCW (7 November 2019): 1–33. https://doi.org/10.1145/3359246.
- Scheuerman, Morgan Klaus, Kandrea Wade, Caitlin Lustig, and Jed R Brubaker. 'How We've Taught Algorithms to See Identity: Constructing Race and Gender in Image Databases for Facial Analysis'. *Proc. ACM Hum.-Comput. Interact.* 4, no. CSCW1 (2020). https://doi.org/10.1145/3392866.
- Schiek, Dagmar. 'Age Discrimination before the ECJ - Conceptual and Theoretical Issues'. *Common Market Law Review*, no. 48 (2011): 777–99.
- ———. 'Broadening the Scope and the Norms of EU Gender Equality Law: Towards a Multidimensional Conception of Equality Law'. *Maastricht Journal of European and Comparative Law* 12, no. 4 (December 2005): 427–66. https://doi.org/10.1177/1023263X0501200405.
- ———. 'From European Union Non-Discrimination Law towards Multidimensional Equality Law for Europe'. In *European Union Non-Discrimination Law: Comparative Perspectives on Multidimensional Equality Law*, 3–27. Routledge-Cavendish Taylor & Francis Group, 2008. https://doi.org/10.4324/9780203892626.
- ———. 'Indirect Discrimination'. In *Materials, Cases and Text on National, Supranational and International Non-Discrimination Law*, edited by Dagmar Schiek, Lisa Waddington, and Mark Bell, 323–75. IUS COMMUNE Casebooks on the Common Law of Europe. Oxford: Hart Publishing, 2007.
- ———. 'On Uses, Mis-Uses and Non-Uses of Intersectionality before the Court of Justice (EU)'. *International Journal of Discrimination and the Law* 18, no. 2–3 (1 June 2018): 82–103. https://doi.org/10.1177/1358229118799232.
- Schuppert, Fabian. *Freedom, Recognition and Non-Domination*. Vol. 12. Studies in Global Justice. Dordrecht: Springer Netherlands, 2014. https://doi.org/10.1007/978-94-007-6806-2.
- ———. 'Non-Domination, Non-Alienation and Social Equality: Towards a Republican Understanding of Equality'. *Critical Review of International Social and Political Philosophy* 18, no. 4 (4 July 2015): 440–55. https://doi.org/10.1080/13698230.2015.1033863.

- ———. ‘Suffering from Social Inequality: Normative Implications of Empirical Research on the Effects of Inequality’. *Philosophical Topics* 40, no. 1 (2012): 97–115.
- Segall, Shlomi. *Why Inequality Matters: Luck Egalitarianism, Its Meaning and Value*. Cambridge: Cambridge University Press, 2016. https://doi.org/10.1017/CBO9781316416969.
- Selbst, Andrew D., and Julia Powles. ‘Meaningful Information and the Right to Explanation’. *International Data Privacy Law* 7, no. 4 (1 November 2017): 233–42. https://doi.org/10.1093/idpl/ipx022.
- Sen, Amartya. ‘Equality of What?’, 26. Stanford University, 1979.
- ———. ‘Essays and Reviews’. In *1959-2002*, 332–38. Princeton University Press, 2014. https://doi.org/10.1515/9781400848393-065.
- ———. ‘One. Merit and Justice’. In *One. Merit and Justice*, 5–16. Princeton University Press, 2018. https://doi.org/10.1515/9780691190334-003.
- ———. *The Idea of Justice*. Belknap Press: Harvard University Press, 2009.
- Sher, George. ‘Why We Are Moral Equals’. In *Do All Persons Have Equal Moral Worth?* Oxford: Oxford University Press, 2014. https://doi.org/10.1093/acprof:oso/9780198719502.003.0002.
- Sherman, Stephen J., Jeffrey W. Sherman, Elise J. Percy, and Courtney K. Soderberg. *Stereotype Development and Formation*. Oxford University Press, 2013. https://doi.org/10.1093/oxfordhb/9780199730018.013.0027.
- Shin, Patrick S. ‘Is There a Unitary Concept of Discrimination?’ In *Philosophical Foundations of Discrimination Law*. Oxford: Oxford University Press, 2013. https://doi.org/10.1093/acprof:oso/9780199664313.003.0009.
- Siapka, Anastasia. ‘Show Me How You Write and I’ll Tell You Who You Are’. CITIP blog, 24 February 2022. https://www.law.kuleuven.be/citip/blog/show-me-how-you-write-and-ill-tell-you-who-you-are/.
- Simonite, Tom. ‘A Health Care Algorithm Offered Less Care to Black Patients’. *Wired*. Accessed 26 April 2022. https://www.wired.com/story/how-algorithm-favored-whites-over-blacks-health-care/.
- ———. ‘Meet the Secret Algorithm That’s Keeping Students Out of College’. *Wired*, 2020. https://www.wired.com/story/algorithm-set-students-grades-altered-futures/.
- Solanke, Iyiola. *Discrimination as Stigma: A Theory of Anti-Discrimination Law.* Oxford, England ; Portland, Oregon: Hart Publishing, 2017.
- Solove, Daniel J. ‘Nothing to Hide: The False Tradeoff Between Privacy and Security (Full Text of Book)’. SSRN Scholarly Paper. Rochester, NY, 31 May 2011. https://papers.ssrn.com/abstract=3976770.
- Solutions, Stephen F. DeAngelis, Enterra. ‘Artificial Intelligence: How Algorithms Make Systems Smart’. *Wired*, 5 September 2014. https://www.wired.com/insights/2014/09/artificial-intelligence-algorithms-2/.
- Somek, Alexander. ‘Moral Re-Engineering: The Neoliberal Predicament’. In *Engineering Equality*. Oxford: Oxford University Press, 2011. https://doi.org/10.1093/acprof:oso/9780199693375.003.0008.
- Sosa, Lorena, and Alexandra Timmer. ‘Conceptual Underpinnings of the EU’s Commitment to Human Rights’. In *The European Union and Human Rights*, by Lorena Sosa and Alexandra Timmer, 62–78. Oxford University Press, 2020. https://doi.org/10.1093/oso/9780198814191.003.0004.
- Sottiaux, Stefan. ‘Het Gelijkheidsbeginsel : Langs Oude Paden En Nieuwe Wegen (Artikel, 2008) [WorldCat.Org]’. *Rechtskundig Weekblad* 72, no. 17 (27 December 2008): 690–702.
- Sottiaux, Stefan, and Jogchum Vrielink. ‘Verschillende Gelijkheden? Recente Europese En Nationale Rechtspraak over Discriminatie’. In *Publiekrecht*, edited by Andre Alen, Gunther Maes, and Stefan Sottiaux, 23–67. Themis Cahier. Brugge: Die Keure, 2012.
- Spinoy, Marie. ‘Discriminatie in Het Gelijkekansendecreet? Reflecties Bij GwH Nr. 110/2019’. *Tijdschrift Voor Bestuurswetenschappen En Publiekrecht*, no. 7 (2020): 317–42.
- Steinhoff, Uwe, ed. *Do All Persons Have Equal Moral Worth? On ‘basic Equality’ and Equal Respect and Concern*. First edition. Oxford: Oxford University Press, 2015.

- Sterba, James P. 'Recent Work on Alternative Conceptions of Justice'. *American Philosophical Quarterly* 23, no. 1 (1986): 1–22.
- Stewart, Heather, Sally Weale, and Kate Proctor. 'Ofqual Chief to Face MPs over Exams Fiasco and Botched Algorithm Grading'. *The Guardian*, 20 August 2020, sec. Education. https://www.theguardian.com/education/2020/aug/20/ofqual-chief-to-face-mps-over-exams-fiasco-and-botched-algorithm-grading-system.
- Stewart, Shannon L., Angela Celebre, John P. Hirdes, and Jeffrey W. Poss. 'Risk of Injury to Others: The Development of an Algorithm to Identify Children and Youth at High-Risk of Aggressive Behaviours'. *Frontiers in Psychiatry* 12 (3 January 2022): 750625. https://doi.org/10.3389/fpsyt.2021.750625.
- Store, ProPublica Data. 'COMPAS Recidivism Risk Score Data and Analysis'. Text/html. ProPublica Data Store. ProPublica Data Store, 2 May 2016. Https://www.propublica.org/datastore/. https://www.propublica.org/datastore/dataset/compas-recidivism-risk-score-data-and-analysis.
- Struys, Bruno. 'Plots stond Bilal op een lijst van de veiligheidsdiensten: "Ik werd bang van mezelf toen ik het rapport las"'. De Morgen, 29 September 2022. https://www.demorgen.be/gs-b11dfd65.
- Surden, Harry. 'Artificial Intelligence and Law: An Overview'. *Georgia State University Law Review* 35, no. 4 (1 June 2019). https://readingroom.law.gsu.edu/gsulr/vol35/iss4/8.
- Sweeney, Latanya. 'Discrimination in Online Ad Delivery', 29 January 2013. https://arxiv.org/ftp/arxiv/papers/1301/1301.6822.pdf.
- Tajfel, Henri. 'Stereotypes'. *Race* 5, no. 2 (1963): 3–14. https://doi.org/10.1177/030639686300500201.
- Tajfel, Henri, and Michael Billic. 'Familiarity and Categorization in Intergroup Behavior'. *Journal of Experimental Social Psychology* 10, no. 2 (1 March 1974): 159–70. https://doi.org/10.1016/0022-1031(74)90064-X.
- Taylor, Linnet, Luciano Floridi, and Bart van der Sloot, eds. *Group Privacy*. Cham: Springer International Publishing, 2017. https://doi.org/10.1007/978-3-319-46608-8.
- Datatilsynet. 'The Norwegian DPA Intends to Order Rectification of IB Grades', 2020. https://www.datatilsynet.no/en/news/2020/the-norwegian-dpa-intends-to-order-rectification-of-ib-grades/.
- Datatilsynet. 'The Norwegian DPA Intends to Order Rectification of IB Grades'. Accessed 18 October 2020. https://www.datatilsynet.no/en/news/2020/the-norwegian-dpa-intends-to-order-rectification-of-ib-grades/.
- Theil, Stefan. 'Is the "Living Instrument" Approach of the European Court of Human Rights Compatible with the ECHR and International Law?', September 2017. https://doi.org/10.17863/CAM.8478.
- Thompson, Michael J. 'Reconstructing Republican Freedom: A Critique of the Neo-Republican Concept of Freedom as Non-Domination'. *Philosophy & Social Criticism* 39, no. 3 (March 2013): 277–98. https://doi.org/10.1177/0191453712473081.
- Thompson, Michael J. 'The Two Faces of Domination in Republican Political Theory'. *European Journal of Political Theory* 17, no. 1 (January 2018): 44–64. https://doi.org/10.1177/1474885115580352.
- Timmer, A. 'Toward an Anti-Stereotyping Approach for the European Court of Human Rights'. *Human Rights Law Review* 11, no. 4 (1 December 2011): 707–38. https://doi.org/10.1093/hrlr/ngr036.
- Timmer, Alexandra. 'A Quiet Revolution: Vulnerability in the European Court of Human Rights'. In *Vulnerability: Reflections on a New Ethical Foundation for Law and Politics*, edited by Martha Fineman and Anna Grear, 147–70. Farnham: Ashgate, 2013. https://lib.ugent.be/en/catalog/pug01:4235269.
- ———. 'Gender Stereotyping in the Case Law of the EU Court of Justice', no. 1 (2016): 10.
- ———. 'Judging Stereotypes: What the European Court of Human Rights Can Borrow from American and Canadian Equal Protection Law'. *The American Journal of Comparative Law* 63, no. 1 (1 January 2015): 239–84. https://doi.org/10.5131/AJCL.2015.0007.

- ———. *Strengthening the Equality Analysis of the European Court of Human Rights: The Potential of the Concepts of Stereotyping and Vulnerability.*, 2014.
- Timmer, Alexandra, and Lourdes Peroni. 'Bah v UK: On Immigration, Discrimination and Worrisome Reasoning'. *Strasbourg Observers* (blog), 12 October 2011. https://strasbourgobservers.com/2011/10/12/bah-v-uk-on-immigration-discrimination-and-worrisome-reasoning/.
- Tobler, Christa. 'Equality and Non-Discrimination under the ECHR and EU Law', 2014, 42.
- ———. *Indirect Discrimination: A Case Study into the Development of the Legal Concept of Indirect Discrimination under EC Law*. Social Europe Series, v. 10. Antwerpen: Intersentia, 2005.
- ———. 'Limits and Potential of the Concept of Indirect Discrimination'. European Network of Legal Experts in the non-discrimination field, 2008.
- Toh. 'Empowering Workers Through Digital Rights'. *Digital Freedom Fund* (blog). Accessed 27 September 2022. https://digitalfreedomfund.org/empowering-workers-through-digital-rights/.
- Tridimas, Takis. *The General Principles of EU Law*. 2nd ed. New York ; Oxford: Oxford University Press, 2006.
- Tzanou, Maria. 'The Future of EU Data Privacy Law: Towards a More Egalitarian Data Privacy Special Issue: Future-Mapping the Directions of EU Law: How Do We Predict the Future of EU Law?' *Journal of International and Comparative Law* 7, no. 2 (2020): 449–70.
- Tzanou, Maria, and Spyridoula Karyda. 'Privacy International and Quadrature Du Net: One Step Forward Two Steps Back in the Data Retention Saga?' SSRN Scholarly Paper. Rochester, NY, 11 July 2021. https://doi.org/10.2139/ssrn.3970756.
- ———. 'Privacy International and Quadrature Du Net: One Step Forward Two Steps Back in the Data Retention Saga?' *EUROPEAN PUBLIC LAW*, n.d., 32.
- Vallentyne, Peter. 'Brute Luck, Option Luck, and Equalityof Initial Opportunities'. *Ethics* 112, no. 3 (April 2002): 529–57. https://doi.org/10.1086/339275.
- ———. 'Discrimination and Rights'. In *The Routledge Handbook of the Ethics of Discrimination*, edited by Kasper Lippert-Rasmussen, 1st ed., 132–39. New York : Routledge, 2017. | Series: Routledge handbooks in applied ethics: Routledge, 2017. https://doi.org/10.4324/9781315681634-13.
- Vallor, Shannon. *Technology and the Virtues: A Philosophical Guide to a Future Worth Wanting*. New York: Oxford University Press, 2016. https://doi.org/10.1093/acprof:oso/9780190498511.001.0001.
- van den Brink, Marjolein and Peter Dunne. 'Trans and Intersex Equality Rights in Europe: A Comparative Analysis.' LU: European Commisson, 2018. https://data.europa.eu/doi/10.2838/75428.
- van den Hoven, Jeroen, and Emma Rooksby. 'Distributive Justice and the Value of Information: A (Broadly) Rawlsian Approach'. In *Information Technology and Moral Philosophy*, edited by Jeroen van den Hoven and John Weckert, 376–96. Cambridge Studies in Philosophy and Public Policy. Cambridge: Cambridge University Press, 2008. https://doi.org/10.1017/CBO9780511498725.019.
- van der Burg, Wibren. 'The Expressive and Communicative Functions of Law, Especially with Regard to Moral Issues'. *Law and Philosophy* 20, no. 1 (1 January 2001): 31–59. https://doi.org/10.1023/A:1006492921790.
- ———. 'The Morality of Aspiration: A Neglected Dimension of Law and Morality'. In *Rediscovering Fuller: Essays on Implicit Law and Institutional Design*, edited by Willem Witteveen J. and Wibren van der Burg, 169–92. Amsterdam: Amsterdam University Press, 2009.
- Van Drooghenbroeck, Sébastien, and Cecilia Rizcallah. 'The ECHR and the Essence of Fundamental Rights: Searching for Sugar in Hot Milk?' *German Law Journal* 20, no. 6 (September 2019): 904–23. https://doi.org/10.1017/glj.2019.68.

- Van Mechelen, Maarten, Jan Derboven, Ann Laenen, Bert Willems, David Geerts, and Vero Vanden Abeele. 'The GLID Method: Moving from Design Features to Underlying Values in Co-Design'. *International Journal of Human Computer Studies* 97 (2017): 116–28. https://doi.org/10.1016/j.ijhcs.2016.09.005.
- Van Mechelen, Maarten, Ann Laenen, Bieke Zaman, Bert Willems, and Vero Vanden Abeele. 'Collaborative Design Thinking (CoDeT): A Co-Design Approach for High Child-to-Adult Ratios'. *International Journal of Human-Computer Studies* 130 (October 2019): 179–95. https://doi.org/10.1016/j.ijhcs.2019.06.013.
- Veale, Michael, and Reuben Binns. 'Fairer Machine Learning in the Real World: Mitigating Discrimination without Collecting Sensitive Data'. *Big Data & Society* 4, no. 2 (December 2017): 205395171774353. https://doi.org/10.1177/2053951717743530.
- Veale, Michael, and Irina Brass. 'Administration by Algorithm?: Public Management Meets Public Sector Machine Learning'. In *Algorithmic Regulation*, edited by Karen Yeung and Martin Lodge, 0. Oxford University Press, 2019. https://doi.org/10.1093/oso/9780198838494.003.0006.
- Veale, Michael, Max Van Kleek, and Reuben Binns. 'Fairness and Accountability Design Needs for Algorithmic Support in High-Stakes Public Sector Decision-Making'. *ArXiv:1802.01029 [Cs]*, 2018, 1–14. https://doi.org/10.1145/3173574.3174014.
- Vedder, Anton. 'KDD: The Challenge to Individualism'. *Ethics and Information Technology* 1, no. 4 (1999): 275–81.
- ———. 'Medical Data, New Information Technologies and the Need for Normative Principles Other than Privacy Rules'. *Law and Medicine. M. Freeman and A. Lewis (Eds.),(Series Current Legal Issues)*, 2000, 441–59.
- Vedder, Anton. 'WHY DATA PROTECTION AND TRANSPARENCY ARE NOT ENOUGH WHEN FACING SOCIAL PROBLEMS OF MACHINE LEARNING IN A BIG DATA CONTEXT'. In *BEING PROFILED*, edited by Emre Bayamlioglu, Irina Baraliuc, Liisa Janssens, and Mireille Hildebrandt, 42–45. COGITAS ERGO SUM: 10 Years of Profiling the European Citizen. Amsterdam University Press, 2018. https://doi.org/10.2307/j.ctvhrd092.10.
- Vedder, Anton H. 'KDD, Privacy, Individuality, and Fairness'. *Readings in CyberEthics, 2nd Edition. Jones and Bartlett, Sudbury, MA*, 2004, 462–70.
- Vedder, Anton, and Laurens Naudts. 'Accountability for the Use of Algorithms in a Big Data Environment'. *International Review of Law, Computers & Technology* 31, no. 2 (4 May 2017): 206–24. https://doi.org/10.1080/13600869.2017.1298547.
- Verbeek, Peter-Paul. 'Materializing Morality: Design Ethics and Technological Mediation'. *Science, Technology, & Human Values* 31, no. 3 (May 2006): 361–80. https://doi.org/10.1177/0162243905285847.
- Verroios, Vasilis, Panagiotis Papadimitriou, Ramesh Johari, and Hector Garcia-Molina. 'Client Clustering for Hiring Modeling in Work Marketplaces'. In *Proceedings of the 21th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 2187–96. KDD '15. New York, NY, USA: Association for Computing Machinery, 2015. https://doi.org/10.1145/2783258.2788589.
- Vervloesem, Koen. 'How Dutch Activists Got an Invasive Fraud Detection Algorithm Banned'. AlgorithmWatch, 2020. https://algorithmwatch.org/en/story/syri-netherlands-algorithm/.
- Vetzo, Max, J. H. Gerards, and Remco Nehmelman. *Algoritmes en grondrechten*. Montaigne Centrum 10. Den Haag: Boom juridisch, 2018.
- Vickers, Lucy. 'ECJ Headscarf Series (2): The Role of Choice; and the Margin of Appreciation'. *Strasbourg Observers* (blog), 8 September 2016. https://strasbourg.weichie.dev/2016/09/08/blog-series-the-role-of-choice-and-the-margin-of-appreciation/.
- Viljoen, Salomé. 'A Relational Theory of Data Governance'. *The Yale Law Journal*, 2021, 82.
- Viljoen, Salome. 'Democratic Data: A Relational Theory For Data Governance'. SSRN Scholarly Paper. Rochester, NY: Social Science Research Network, 11 November 2020. https://doi.org/10.2139/ssrn.3727562.

- Virginia Mantouvalou. 'Work and Private Life: Sidabras and Dziautas v Lithuania'. *European Law Review*, no. 30 (10 August 2005): 573–85.
- Visser, Froukje Sleeswijk, Pieter Jan Stappers, Remko van der Lugt, and Elizabeth B.-N. Sanders. 'Contextmapping: Experiences from Practice'. *CoDesign* 1, no. 2 (1 April 2005): 119–49. https://doi.org/10.1080/15710880500135987.
- Vogiatzoglou, Plixavra. 'Mass Surveillance, Predictive Policing and the Implementation of the CJEU and ECtHR Requirement of Objectivity'. *European Journal of Law and Technology* 10, no. 1 (16 May 2019). https://ejlt.org/index.php/ejlt/article/view/669.
- Voigt, Kristin. 'Distributive Equality, Relational Equality and Preferences about Higher Education'. *Theory and Research in Education* 15, no. 2 (1 July 2017): 109–28. https://doi.org/10.1177/1477878517708116.
- ———. 'Relational Equality and the Expressive Dimension of State Action'. *Social Theory and Practice* 44, no. 3 (2018): 437–67.
- ———. 'The Harshness Objection: Is Luck Egalitarianism Too Harsh on the Victims of Option Luck?' *Ethical Theory and Moral Practice* 10, no. 4 (1 August 2007): 389–407. https://doi.org/10.1007/s10677-006-9060-4.
- Wachter, Sandra. 'Affinity Profiling and Discrimination by Association in Online Behavioral Advertising'. *Berkeley Technology Law Journal* 35, no. 2 (2020): 367–430.
- ———. 'Affinity Profiling and Discrimination By Association in Online Behavioral Advertising', 2021. https://doi.org/10.15779/Z38JS9H82M.
- ———. 'Affinity Profiling and Discrimination by Association in Online Behavioural Advertising'. SSRN Scholarly Paper. Rochester, NY: Social Science Research Network, 15 May 2019. https://papers.ssrn.com/abstract=3388639.
- ———. 'The Theory of Artificial Immutability: Protecting Algorithmic Groups under Anti-Discrimination Law'. *SSRN Electronic Journal*, 2022. https://doi.org/10.2139/ssrn.4099100.
- Wachter, Sandra, and Brent Mittelstadt. 'A Right to Reasonable Inferences: Re-Thinking Data Protection Law in the Age of Big Data and AI' 2019, no. 2 (2019): 494–620. https://doi.org/10.7916/d8-g10s-ka92.
- Wachter, Sandra, Brent Mittelstadt, and Luciano Floridi. 'Why a Right to Explanation of Automated Decision-Making Does Not Exist in the General Data Protection Regulation'. *International Data Privacy Law* 7, no. 2 (2017): 76–99.
- Wachter, Sandra, Brent Mittelstadt, and Chris Russell. 'Bias Preservation in Machine Learning: The Legality of Fairness Metrics Under EU Non-Discrimination Law'. SSRN Scholarly Paper. Rochester, NY: Social Science Research Network, 15 January 2021. https://doi.org/10.2139/ssrn.3792772.
- ———. 'Counterfactual Explanations without Opening the Black Box: Automated Decisions and the GDPR', 2017.
- ———. 'Why Fairness Cannot Be Automated: Bridging the Gap Between EU Non-Discrimination Law and AI'. SSRN Scholarly Paper. Rochester, NY: Social Science Research Network, 3 March 2020. https://doi.org/10.2139/ssrn.3547922.
- Waddington, Lisa. 'Saying All the Right Things and Still Getting It Wrong: The Court of Justice's Definition of Disability and Non-Discrimination Law'. *Maastricht Journal of European and Comparative Law* 22, no. 4 (August 2015): 576–91. https://doi.org/10.1177/1023263X1502200406.
- Wagner, Ben, Paola Lopez, Florian Cech, Gabriel Grill, and Marie-Therese Sekwenz. 'Der AMS-Algorithmus.' *zeitschrift für kritik & recht gesellschaft* 2020, no. 2 (2020): 191–202. https://doi.org/10.33196/juridikum202002019101.
- Waldron, Jeremy. *One Another's Equals: The Basis of Human Equality*. Cambridge, Massachusetts ; London, England: Belknap Press: Harvard University Press, 2017.
- Walzer, Michael. *Spheres of Justice: A Defense of Pluralism and Equality*. Basic Books, Inc., 1983.

- Wang, Yilun, and Michal Kosinski. 'Deep Neural Networks Are More Accurate than Humans at Detecting Sexual Orientation from Facial Images.', 15 February 2017. https://doi.org/10.17605/OSF.IO/ZN79K.
- Ward, Angela. 'The Impact of the EU Charter of Fundamental Rights on Anti-Discrimination Law: More a Whimper than a Bang?' *Cambridge Yearbook of European Legal Studies* 20 (December 2018): 32–60. https://doi.org/10.1017/cel.2018.11.
- Weinberger, Lisa Maria. 'Kurt v. Austria: A Missed Chance to Tackle Intersectional Discrimination and Gender-Based Stereotyping in Domestic Violence Cases'. *Strasbourg Observers* (blog), 8 August 2021. https://strasbourgobservers.com/2021/08/18/kurt-v-austria-a-missed-chance-to-tackle-intersectional-discrimination-and-gender-based-stereotyping-in-domestic-violence-cases/.
- Wenar, Leif. 'John Rawls'. In *The Stanford Encyclopedia of Philosophy*, edited by Edward N. Zalta, Spring 2017. Metaphysics Research Lab, Stanford University, 2017. https://plato.stanford.edu/archives/spr2017/entries/rawls/.
- Westen, P. *Speaking of Equality: An Analysis of the Rhetorical Force of 'Equality' in Moral and Legal Discourse*. *Speaking of Equality*. Princeton University Press, 2014. https://doi.org/10.1515/9781400861484.
- Westen, Peter. 'The Empty Idea of Equality'. *Harvard Law Review* 95, no. 3 (1982): 537–96.
- Williams, Patrick, and Eric Kind. 'Data-Driven Policing: The Hardwiring of Discriminatory Policing Practices across Europe.' European Network Against Racism, 2019.
- Wit, T.W.A. de. 'Wanneer Iedereen Aan Zijn Eigen Waarheid de Voorkeur Geeft. Kan Politiek Wel Zonder Common Ground?' In *Onenigheid En Gemeenschap. Basisboek Politieke Filosofie*, edited by M. Terpstra, 528–48. Amsterdam: Boom, 2012.
- Witte, Bruno de. 'From a "Common Principle of Equality" to "European Antidiscrimination Law"'. *American Behavioral Scientist* 53, no. 12 (August 2010): 1715–30. https://doi.org/10.1177/0002764210368093.
- Wolff, Jonathan, and Avner de-Shalit. *Disadvantage*. *Disadvantage*. Oxford University Press, 2007. https://www.oxfordscholarship.com/view/10.1093/acprof:oso/9780199278268.001.0001/acprof-9780199278268.
- ———. 'On Fertile Functionings: A Response to Martha Nussbaum'. *Journal of Human Development and Capabilities* 14, no. 1 (February 2013): 161–65. https://doi.org/10.1080/19452829.2013.762177.
- Woods, Dorian R., Yvonne Benschop, and Marieke van den Brink. 'What Is Intersectional Equality? A Definition and Goal of Equality for Organizations'. *Gender, Work & Organization* 29, no. 1 (2022): 92–109. https://doi.org/10.1111/gwao.12760.
- Writer, Christina Pazzanese Harvard Staff. 'How Unjust Police Killings Damage the Mental Health of Black Americans'. *Harvard Gazette* (blog), 13 May 2021. https://news.harvard.edu/gazette/story/2021/05/how-unjust-police-killings-damage-the-mental-health-of-black-americans/.
- Xenidis, Raphaele. 'Multiple Discrimination in EU Anti-Discrimination Law: Towards Redressing Complex Inequality?' *EU Anti-Discrimination Law Beyond Gender*, 15 November 2018. https://www.research.ed.ac.uk/en/publications/multiple-discrimination-in-eu-anti-discrimination-law-towards-red.
- ———. 'Shaking the Normative Foundations of EU Equality Law: Evolution and Hierarchy between Market Integration and Human Rights Rationales', 2017. https://www.research.ed.ac.uk/en/publications/shaking-the-normative-foundations-of-eu-equality-law-evolution-an.
- Xenidis, Raphaële. 'The Polysemy of Anti-Discrimination Law: The Interpretation Architecture of the Framework Employment Directive at the Court of Justice.' *Common Market Law Review* 58 (2021): 1649–96.
- ———. 'Transforming EU Equality Law? On Disruptive Narratives and False Dichotomies'. *Yearbook of European Law* 38 (1 January 2019): e2–57. https://doi.org/10.1093/yel/yey005.

- ———. ‘Tuning EU Equality Law to Algorithmic Discrimination: Three Pathways to Resilience’. *Maastricht Journal of European and Comparative Law* 27, no. 6 (December 2020): 736–58. https://doi.org/10.1177/1023263X20982173.
- Xenidis, Raphaële, and Linda Senden. ‘EU Non-Discrimination Law in the Era of Artificial Intelligence : Mapping the Challenges of Algorithmic Discrimination’. In *General Principles of EU Law and the EU Digital Order*, edited by Ulf Bernitz, Xavier Groussot, Jaan Paju, and Sybe A. De Vries, 151–82. Kluwer Law International, 2020. https://cadmus.eui.eu//handle/1814/65845.
- ‘Xenophobic Machines: Discrimination through Unregulated Use of Algorithms in the Dutch Childcare Benefits Scandal.’ Amnesty International, October 2021. https://www.amnesty.nl/content/uploads/2021/10/20211014_FINAL_Xenophobic-Machines.pdf?x42580.
- Yeung, Karen. ‘“Hypernudge”: Big Data as a Mode of Regulation by Design’. *Information, Communication & Society* 20, no. 1 (2 January 2017): 118–36. https://doi.org/10.1080/1369118X.2016.1186713.
- Yeung, Karen, and Martin Lodge, eds. *Algorithmic Regulation*. Oxford, New York: Oxford University Press, 2019.
- Young, Iris Marion. ‘Equality of Whom? Social Groups and Judgments of Injustice’. *Journal of Political Philosophy* 9, no. 1 (1 March 2001): 1–18. https://doi.org/10.1111/1467-9760.00115.
- ———. *Inclusion and Democracy*. Oxford University Press, 2002. https://doi.org/10.1093/0198297556.001.0001.
- ———. *Justice and the Politics of Difference*. Princeton, New Yersey: Princeton University Press, 1990.
- Zarsky, Tal. ‘Transparent Predictions’. SSRN Scholarly Paper. Rochester, NY: Social Science Research Network, 10 September 2013. https://papers.ssrn.com/abstract=2324240.
- Zarsky, Tal Z. ‘An Analytic Challenge: Discrimination Theory in the Age of Predictive Analytics Cross-Cutting Issues of Law’. *I/S: A Journal of Law and Policy for the Information Society*, no. 1 (2018 2017): 11–36.
- ———. ‘Understanding Discrimination in the Scored Society’. *Wash. L. Rev.* 89, no. 4 (2014): 1375–1412.
- Zimmermann, Annette. ‘The Power of Choosing Not to Build: Justice, Non-Deployment, and the Purpose of AI Optimization’, 2021.
- Zuboff, Shoshana. *The Age of Surveillance Capitalism: The Fight for the Future at the New Frontier of Power*. London: Profile Books, 2019.
- Zuiderveen Borgesius, Frederik, and Joost Poort. ‘Online Price Discrimination and EU Data Privacy Law’. *Journal of Consumer Policy* 40, no. 3 (1 September 2017): 347–66. https://doi.org/10.1007/s10603-017-9354-z.