UNIVERSITY OF AMSTERDAM

UvA-DARE (Digital Academic Repository)

Designing for justice in freelancing

Testing platform interventions to minimise discrimination in online labour markets

Brooke, Siân; Rao, Aliya Hamid

DOI
10.1177/20539517241232631

Publication date
2024

Document Version
Final published version

Published in
Big Data and Society

License
CC BY

Link to publication

Citation for published version (APA):
Brooke, S., & Rao, A. H. (2024). Designing for justice in freelancing: Testing platform interventions to minimise discrimination in online labour markets. *Big Data and Society*, *11*(1). https://doi.org/10.1177/20539517241232631

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:17 Oct 2026

Check for updates

Original Research Article

BIG DATA & SOCIETY

Big Data & Society
January–March: 1–19
© The Author(s) 2024
Article reuse guidelines:
sagepub.com/journals-permissions
DOI: 10.1177/20539517241232631
journals.sagepub.com/home/bds
Sage

Designing for justice in freelancing: Testing platform interventions to minimise discrimination in online labour markets

Siân Brooke[1] **and Aliya Hamid Rao**[1]

Abstract

Online labour markets (OLMs) are a vital source of income for globally diverse and dispersed freelancers. Despite their promise of neutrality, OLMs are known to perpetuate hiring discrimination, vested in how OLMs are designed and what kinds of interactions they enable between freelancers and hirers. In this study, we go beyond understanding mechanisms of hiring discrimination in OLMs, to identifying platform design features that can minimise hiring discrimination. To do so, we draw on a methodology guided by the design justice ethos. Drawing on a survey on UK-based freelancers and interviews with a purposefully drawn sub-sample, we collaboratively identify five platform design interventions to minimise hiring discrimination in OLMs: community composition, identity-signalling flairs, text only reviews, union membership, and an antidiscrimination prompt. The core of our study is an innovative experiment conducted on a purpose-built, mock OLM, *Mock-Freelancer.com*. On this mock OLM, we experimentally test mechanisms of discrimination, including how these mechanisms fare under the five altered platform design interventions through a discrete-choice experiment. We find that both community and flairs were important in encouraging the hiring of women and non-White freelancers. We also establish that anonymity universally disadvantages freelancers. We conclude with recommendations to design OLMs that minimise labour market discrimination.

Keywords

Future of work, gender, design justice, ethnicity, platforms, affordances

Introduction

Online labour markets (OLMs) have witnessed rapid growth in recent years, expanding opportunities for freelancing (Stephany et al., 2021). Wood et al. (2019) estimate that 163 million freelancers are registered on OLMs worldwide, with 4.2 million in the UK reporting freelancing as their main job.[1] Data suggest that a third of freelancers come from India, especially for software development projects (Stephany et al., 2021). Freelancers on OLMs are occupationally heterogeneous and globally diverse (Leung, 2018). Women comprise about 40% of online freelancers (Stephany et al., 2021). Freelancing jobs range from manual labour such as carpentry and plumbing, domestic labour to software development, writing, and creative and multimedia work (Kässi and Lehdonvirta, 2018). Part of OLMs' appeal to employers is the ability to hire on a short-term basis for specific needs and to do so by accessing a skilled global labour force. Following platform economy scholars, such as Kässi and Lehdonvirta, (2018) and Stephany et al., (2021), we conceptualise these platforms as marketplaces to highlight our focus on the labour market (i.e. hiring) aspects of platforms.

Given the ethnic and gender diversity that encompasses online freelancing, attention needs to be paid to how job opportunities are distributed across OLM freelancers. Who are the people who get selected for jobs? Discrimination in labour markets works through mechanisms which often sort individuals into occupations deemed most suitable given an individual's status characteristics, such as ethnicity and gender (Hong, 2016; Ridgeway and Cornell, 2006). Labour markets with high-status and high-paid jobs typically tend to favour White men, who benefit from stereotypes of expertise and

[1]Department of Methodology, London School of Economics and Political Science, London, UK

Corresponding author:
Siân Brooke, Department of Methodology, London School of Economics and Political Science, London, UK.
Email: s.j.brooke@lse.ac.uk

Creative Commons CC BY: This article is distributed under the terms of the Creative Commons Attribution 4.0 License (https://creativecommons.org/licenses/by/4.0/) which permits any use, reproduction and distribution of the work without further permission provided the original work is attributed as specified on the SAGE and Open Access page (https://us.sagepub.com/en-us/nam/open-access-at-sage).

assumed leadership ability. Women are often perceived by hirers as 'naturally' less skilled in technical professions and less committed to competitive workplaces (Brooke, 2021; Rivera and Tilsik, 2016). In lower wage labour markets, there is also a racialised and gendered preference (Hong, 2016), such as for women from the Global South for sewing and embroidery based on these women's ostensibly 'nimble fingers' and the preference for Asian women for beauty work in the US (Kang, 2010). Different labour markets thus have different mechanisms of discrimination.

In offline labour markets, research has also shown how hiring discrimination can be managed through slight tweaks to the hiring process, for example, by having a blinded orchestra audition (Goldin and Rouse, 2000). On online platforms, scholars have used the conceptual tool of 'affordances' (Davis, 2020) to understand discrimination. Affordances emphasise how an object (i.e. platform) interfaces with its social environment to, at times, produce biased and discriminatory outcomes (Noble, 2018; Schwartz and Neff, 2019). OLMs are conducive for examining how labour market discrimination can be minimised. Platform owners can make platform design decisions that amplify or minimise different biases via the hiring platform. This relative ease of manipulating platform design is partly why OLMs have been framed as spaces where hiring discrimination may be minimised (Stringhi, 2022). An example of a platform design decision is curtailing users' active decision-making in hiring someone in favour of algorithmic 'matching' (Wood and Lehdonvirta, 2021). For example, Uber does not let riders pick from a range of drivers (Levy and Barocas, 2017). Leaving matching to algorithms brings its own host of problems, wherein discrimination may be baked into the very algorithm (Jarrahi et al., 2021; Noble, 2018).

While prior research has successfully identified mechanisms of labour market discrimination, how hiring platforms can be designed to minimise discrimination has received less attention. We follow the lead of scholars seeking to identify and minimise labour market discrimination, by focusing on online freelancing (Correll, 2017; Levy and Barocas, 2017). Researching the functioning of discrimination in freelancing is crucial to ensuring that hiring discrimination is challenged as work moves increasingly online. Building on large-scale statistical analysis of OLMs and qualitative research on the experiences of freelancers (Hannák et al., 2017), we develop an innovative mixed-methods study to test which design interventions can help minimise biases and discrimination in freelancing work adjudicated via OLMs. We ask the following research questions:

RQ1: What forms of hiring discrimination are present in OLMs?
RQ2: How do different platform-based interventions affect hiring discrimination in OLMs?

Our methodology is motivated by design justice (Baumer, 2017; Costanza-Chock, 2020), drawing on a survey ($n = 884$) and interview data ($n = 18$) with freelancers who use OLMs to determine five design interventions that may reduce discrimination in OLMs. In our primary methodological component, we then experimentally model these design interventions on a purpose-built mock online platform (*Mock-Freelancer.com*) to understand which design interventions can most effectively reduce discrimination in OLMs.

Our key contribution is to show design interventions that can, with relative ease, minimise discrimination in OLMs. The study furthers social science research to identify how and when bias and discrimination are most acute and when they can be adjusted to ensure that all freelancers gain access to freelance work through OLMs. This is important because OLMs are a source of income for millions of workers. OLMs also become critical as an income source during economic crises (International Labour Organisation, 2021). Our findings demonstrate evidence of a general preference for South Asian women freelancers, which increases under specific interventions. We find community composition and flairs to be especially useful for minimising discrimination in this study. Anonymity, previously seen as potentially curtailing discrimination in online and offline labour markets (Stefanelli and Lukac, 2020), does not appear to benefit freelancers in this study. Our findings contribute to research that, in addition to identifying mechanisms of labour market discrimination, also provides evidence on designing interventions to reduce discrimination.

Literature review

Mechanisms and conditions of affordances in OLMs

Algorithms and platforms are not neutral, objective entities that inevitably produce equal outcomes (Davis, 2020; Eubanks, 2018; Noble, 2018). Instead, technologies are social. Scholars studying how technologies interface with social environments emphasise how biases can be built into technologies producing discriminatory outcomes (Noble, 2018). Especially salient in this line of conceptualisation is the notion of 'affordances' which takes into account how, beyond the technical capacity of a design object, the use of an object depends substantially on its social environment and conditions (Davis, 2020). Moving beyond a dichotomy of what can or cannot be done with a technology, Davis (2020) develops the 'mechanisms and conditions framework'. The *mechanism* through which an affordance operates is not applied as a binary (does, does not afford) but is a continuum by which technologies *request, demand, encourage, discourage, refuse,* and allow specific kinds of social action. A platform may *request* a user provides a birthdate (i.e. Reddit.com) or it

may *demand,* not allowing the creation of an account without it (Facebook.com). Effectively neutral, *allow* does not pressure users to take one action or another. Davis's (2020) model also addresses the problems of a presumed universal subject; the *conditions* of affordance considers how social outcomes will depend on the type of users and the particularities of users' social contexts. How users interact with technology is impacted by their technological literacy, choice in device, ability to manipulate the device, and the cultural and institutional context in which they are using it (Davis, 2020).

The mechanisms and conditions of affordances framework are useful for understanding how platforms shape labour practices in online work (Bellesia et al., 2023; Schwartz and Neff, 2019; Wood et al., 2019). In her study of the ride-hailing app Uber, Rosenblatt (2019) shows that while Uber promotes driving as characterised by freedom and independence, Uber's platform design disadvantages its workers through how it structures labour and pay. Through the affordances of *request* mechanisms, workers are incentivised to work longer hours with less flexibility. An example of this is 'surge pricing' where the algorithmic assessment of supply and demand temporarily raises fares for a particular geographic location. Through such algorithmic structures, Uber exerts a soft control over its workers, challenging their independence.

Like Rosenblatt (2019), Wood et al.'s (2019) study of online freelancing found that flexibility promised by platforms was the primary driver for freelancers seeking work through OLMs. However, freelancers had to employ algorithmic management techniques through curating platform-based ratings and reputation systems to maintain autonomy in their work by appealing to hirers. How platforms compute the scores, or ratings and rankings that a user has on a particular platform, are often opaque to users. Users thus deploy algorithmic management in varied ways. Some freelancers *perceive* (Davis, 2020) scores as instrumental to their visibility on platforms, but for others, scores become an extension of themselves and they view these scores as an identity statement about their quality of labour (Bellesia et al., 2023). Platform owners appear to encourage this ambiguity, viewing the ambiguity as a mechanism for hirers to ascertain the quality of freelancers (Bellesia et al., 2023). Platform norms are often to give freelancers high scores for their work (Garg and Johari, 2019). Consequently, *rating inflation* poses a challenge to freelancers' careful self-curation on platforms. Indeed, Wagner et al. (2021) highlight that rating inflation results in a prevalence of 'perfect' worker ratings. Given rating inflation, high scores are not necessarily treated by hirers as providing useful information about a freelancer's quality of labour. Hirers thus draw on other characteristics that are visible, or not, on a freelancer's profile.

Affordances have individual and collective implications for structuring interactions between different users. A mechanism that may appear neutral to one group may be discouraging or even hostile to another (Davis, 2020). Schwartz and Neff (2019) further the concept of affordances by arguing how affordances are gendered because they 'enable different users to take different actions based on the gendered social and cultural repertories available to users and technology designers' (Schwartz and Neff, 2019: p.2407). Stringhi (2022) illustrates how the affordances associated with controls over information flows and interactions between different categories of users extend gender affordances to freelancing platforms. To establish reputation and build trust, OLMs encourage the creation of personal profiles that feature first names, encourage an identifying profile picture, and even linking to external social media accounts. Stringhi (2022) shows that this mechanism of affordance triggers unintended consequences, such as facilitating racial discrimination and harassment of freelancers. This personalisation in OLMs emphasises the significance of the identity of the transacting parties, enhancing the probability for gender discrimination (Stringhi, 2022). Yet, technologies, and specifically platforms, can also be thoughtfully designed keeping in mind what kind of affordances each design choice renders possible.

Hiring discrimination in OLMs

Platform design features are key in shaping both the mechanisms and conditions of affordances in discrimination in OLMs. It is important to thus pinpoint precisely which features, when, and for whom, produce discrimination and how such discrimination may be mitigated. Research on discrimination in OLMs is essential, albeit currently nascent, given how OLMs are a vital source of labour for many (Stephany et al., 2021). The comprehensive research on discrimination in traditional labour markets informs us that ethnicity and gender are two status characteristics that tend to confer advantage on White men (Gaddis, 2015; Quadlin, 2018; Rivera and Tilcsik, 2016; Tak et al., 2019). White men benefit from being considered competent and being viewed as having leadership skills and 'natural' ability for some of the highest-paying positions. This translates into hiring, pay, and promotion advantages. In contrast, stereotypes about women, like their suitability for tasks viewed to be masculine, such as coding (Brooke, 2021) and Math, often disadvantage them, especially in elite labour markets (Quadlin, 2018; Rivera and Tilcsik, 2016). Similar mechanisms may be at play in OLMs. OLM employers overwhelming tend to be based in majority White and developed countries of the Global North such as the United States and the United Kingdom (UK), while the highly educated workforce of freelancers they hire from is substantially located in countries of the Global South, such as India (Stephany et al., 2021). This creates a space ripe for racialised and gendered stereotypes to be activated

as White hirers make decisions about hiring from a racially diverse workforce.

The emerging literature on discrimination in OLMs (Galperin, 2021; Hangartner, Kopp and Siegenthaler, 2021; Pallais, 2014) finds that hiring biases advantage women from less developed countries (Chan and Wang, 2018). Biases advantaging women in OLMs tend to be concentrated in jobs that are feminine-typed, such as administration (Cerqua and Urwin, 2018). Women's disadvantage remains for high-value projects that pay more (Cerqua and Urwin, 2018). Women from the Global South have often been viewed as suited for detail-oriented tasks (Kang, 2010). Mechanisms leading to the preference for South Asian women in OLMs may be vested in similar stereotypes about women. Gendered occupational deviation appears to disadvantage men and women in OLMs. In a large-scale study of an online Swiss recruitment platform, Hangartner et al. (2021) track how much time recruiters spend on each profile. Although recruiters spend similar amounts of time on all profiles, they are less likely to contact women in male-dominated fields and men in female-dominated fields, as well as individuals from immigrant and minority ethnic groups. The mechanism for penalty for women in masculine-typed jobs appears to be that employers fall back on gender stereotypes in the absence of information, and these stereotypes reduce women's chances of getting hired (Hong, 2016).

Hirers' decisions could be tethered to offline gendered biases, but another possibility is that hirers' biases are reinforced by each platform. A common feature on OLMs is showing the 'community' of workers who have completed a particular task. A task could be reinforced as being masculine-typed if hirers are shown images of men who have previously completed this or a similar task. But this same assumption of a task as masculine-typed could be undermined if hirers are shown a more gender-balanced community of workers as previously completing this task. Indeed, employers appear to update their assumptions about workers as they gain familiarity with different features of a platform (Galperin, 2021). This indicates that OLM employers can be trained to hire in ways that minimise discrimination.

Intervening in discrimination in OLMs

Scholarship from the tradition of human–computer interaction has argued that platforms should challenge implicit processes of discrimination through *design justice* (Baumer, 2017; Costanza-Chock, 2020). Design justice is a field of theory and practice concerned with how the design of objects and systems influences the distribution of risks, harms, and benefits among groups of people (Dekker et al., 2022; Viljoen, Goldenfein and McGuigan, 2021). Focusing on how platforms can be 'designed against discrimination', Levy & Barocas (2017: p.1186) point to 10 design possibilities under three strategies that could mitigate against discrimination: (a) 'setting policies' (with design interventions: engaging under-represented groups in the design process; restructuring community through norms; community guidelines), (b) 'structuring interactions' (prompting and priming; how users learn about each other; what users learn about each other; reputation and ratings), and (c) 'monitoring and evaluating' (reporting and sanctioning; data quality; measurement and detection).

The information required for an online freelancing profile typically consists at a basic level of an image (or profile photo), username, average rating, number of reviews, location, and skills. Prior research indicates that reputation systems on platforms, based on ratings and reviews, are a crucial mechanism in making some freelancers more coveted than others on OLMs. Reputation systems replicate broader cultural prejudices despite ostensibly being founded in quantifiable approaches to fairness and justice (Hannák et al., 2017). Even the architecture of ratings is conducive to amplifying gendered discrimination (Rivera and Tilcsik, 2019). One study shows that when candidates were ranked on a scale of 1– 6 versus 1–10, the gender gap in evaluation widens substantially. In the 1–10 scale, the gender gap in the evaluation was significantly larger, with the mechanism being that participants were wary of ranking women a 10, which they viewed as indicating an overly generous score for women. Evaluators were far more comfortable scoring women at the highest level of 6 on a 1–6 scale because they perceived that this scale did not indicate an unduly inflated evaluation (Rivera and Tilcsik, 2019). Scores and rankings, so central in OLMs, are often a biased metric.

A mechanism of affordance is how platforms can allow businesses to highlight specific aspects of their identity. In 2021, both Instagram and Google added a feature that allowed businesses to add a 'business identity attribute' or flair, such as 'black-owned' or 'woman-owned'. Research has found that labelling restaurants as Black-owned on Yelp.com increased customer engagement and the number of reviews left by White customers (Aneja et al., 2023). Similarly, Babar et al. (2023) found that the woman-owned identity signalling also increases business for profiles on Google Maps. Aneja et al. (2023) suggest that these positive effects are because consumers desire to compensate for the disadvantages and discrimination that women and ethnic minorities face. Research on identity-signalling flairs is restricted to reviews of offline businesses through review platforms and has not included freelancing work to date. While anonymity can protect freelancers from discrimination, there is also evidence that identity signalling too is beneficial.

Because prior research has pointed to how men in feminine-typed tasks and women in masculine-typed tasks are more likely to experience discrimination in OLMs in

terms of being hired, it makes sense to pay attention to this aspect of hiring in OLMs through focusing on the community of workers. Moreover, given that the racial and gender background of an individual matters for hiring in both offline and OLMs, information that indicates this background (i.e. through photograph, name, or an explicit declaration, such as being a minority-owned or woman-owned business) requires attention as well. We draw on previous scholarship on the affordances approach, hiring discrimination, and design justice to first identify specific platform design features most conducive to amplifying hiring discrimination and then to manipulate these features to provide experimental evidence to minimise discrimination in OLMs.

Our key contribution is methodological: we create a mock platform that allows us to experimentally test different platform design choices. Scholars have typically not been privy to the design choices of real OLMs like Freelancer or Upwork. The algorithmic decisions made on those platforms remain a black box. Further, the valuable conceptual work on affordances and design justice is often not empirically tested. This arguably provides platforms with a justification for not pursuing such interventions. Because we created our own mock platform, we can cleanly manipulate the relevant design feature and separate out how and when hiring discrimination is amplified and how it can be minimised through selective and considered manipulation of each design feature. This study extends prior research on discrimination in OLMs which has tended to rely on survey and tracking data (Hangartner et al., 2021). Our study goes beyond describing that discrimination occurs to testing whether, and which, design features can be most effectively manipulated to minimise discrimination. This methodological contribution enables us to contribute to the growing scholarship on minimising labour market discrimination.

Methodology

Our research design consisted of a survey, interviews, and a discrete-choice experiment. Given RQ1, we designed a survey of UK freelancers to elicit their perceptions of hiring discriminations on OLMs. Next, we conducted interviews with a purposefully selected sub-sample of freelancers who participated in our survey. Through these interviews, we probed for platform design changes that freelancers think can minimise hiring discrimination. This aspect of our methodology, treating freelancers as 'peer researchers' who are involved in the platform design, is shaped by design justice (Davis, 2020) and bridges RQ1 and RQ2 by encouraging freelancers to draw on their experiences of discrimination to inform interventions. Our final step, to answer RQ2, was to conduct a discrete-choice experiment to understand which platform design choices can help minimise hiring discrimination.

Survey of freelancers

We developed a survey to collect freelancers' experiences on how they experience discrimination. The survey was completed in July 2022 and yielded 884 valid responses. Respondents were reimbursed £7.50 for their time and to motivate detailed, free-text answers. We contacted freelancers by posting on social media forums for freelancers in the UK, specifically Facebook groups (i.e. Freelancers UK) and Reddit (r/freelanceuk). While non-probability sampling is often the gold standard in survey research, for comparison with existing research, we limited the sample to those who either lived in the UK or had British citizenship (Lehdonvirta et al., 2021). We focused on freelancers living in the UK to ensure internal validity as we are focusing on culturally shared notions of gender and ethnicity. However, the sampling frame does limit the generalisability of the findings outside of the UK. We determined that the sample characteristics were representative of the gender and ethnicity of the wider UK online freelancing population in terms of equivalent stratification identity categories. [I] 55% of the sample was women, 45% men, and <1% identified as non-binary or gender-queer. The mean age was 31 years. 93% of respondents disclosed their ethnicity; 53% were White, 30% identified as from Multiple Ethnic Groups, 6% were South Asian, and 4% were Black. There is likely an economic self-selection effect (Lehdonvirta et al., 2021), where the opportunity cost of participating in the survey is lower for lower-paid freelancers. A summary of the main themes of the survey is in Table 1.

Interviews with freelancers

At the close of the survey, freelancers indicated if they could be contacted for an interview. 39% of freelancers who completed the survey agreed to be contacted for interview. Given our research focus on discrimination in online freelancing, we prioritised interviewing freelancers across a wide range of those who felt they were not discriminated against at all in OLMs to those who felt they were. We created a metric of discrimination frequency by combining and re-coding the frequency of discriminatory experiences from the survey to a numeric scale. In our survey, this corresponded to the following item: *Indicate the frequency that you have experienced the below behaviours while freelancing*. The 26 experiences we used to gauge discrimination included: 'the time you dedicate to a project challenged or questioned by clients', 'You have been called names or insulted by other freelancers', and 'Didn't appear in search results when you should have'. The potential multiple-choice responses and code were 'Never': 0, 'Yes, but not in the past year': 1, 'Yes, once or twice in the past year': 2, 'Yes, many times in the past year': 3. These discriminatory scenarios were informed by

Table 1. Summary of the survey instrument themes and content.

Survey theme	Content / Questions
Demographics	Gender identity; Age; Ethnic group; Disability; Education; Employment; Financial situation
Freelancing Profile	Occupation category; Percentage monthly income; Day/Hourly rate; Experience; Proficiency; Work schedule
Freelancing Strategy	Platform profile; Anonymity; Platform choice; Advertising
Opinions on Freelancing	Positive statements: Important attributes in being hired; Negative statements
Being Paid	Refused work; Any discount; 50% < discount; Free
Discrimination and Hostility	Clients; Colleagues; Platforms; Specific Behaviours; Abuse
Response	Reporting; Advice

existing literature, primarily Hannak (2017), and the experiences described in the Big Freelancer Survey (2022) and International Labour Organisations Flagship Report (2021). We coded these to a numeric scale corresponding to relatively low (0–0.2), medium (0.4–0.6), or high (0.8–1.0) frequency of discrimination. The density of discrimination across all freelancers is shown in Figure 1. We then chose 18 freelancers to interview based on their experiences of discrimination.

We aimed for our interview sample to capture a variation in pay rates, as we discuss later in the results. We also paid attention to interviewing freelancers across salient dimensions of stratification, including gender, ethnicity, and variation in six occupation classes as identified by Kässi and Lehdonvirta (2018). There were three freelancers from each of the occupation classes: (a) Clerical and data entry, (b) Creative and multimedia, (c) Professional services, (d) Sales and marketing support, (e) Software development and technology, and (f) Writing and translation. These occupation classes categorise the breadth of online labour undertaken on freelancing platforms (Stephany et al., 2021). The sample consisted of 10 women and eight men. In terms of ethnicity: five identified as White, four as Black, three as East Asian, and six as South Asian.

We designed our interview protocol to gain deeper insight into what aspects of OLMs freelancers saw as most salient in their experiences of discrimination and what platform design features they saw as beneficial to minimising discrimination (see Table 2). The first author and a graduate research assistant conducted the interviews via online video. Interviews lasted approximately 30–40 min. Interviewees were reimbursed £30.00 each, paid via Amazon voucher. This sum was to appropriately acknowledge interviewees' input and time towards the research goals. Our interview and survey data combined led us to prioritise the interventions in to test in our experiment.

Twelve interviewees highlighted the importance of being a native English speaker for being hired. Freelancers shared that so-called 'proficiency of English', regardless of spoken ability, was used to justify refusing work or firing freelancers from a project. Freelancers referred to English proficiency as a 'proxy' for racism. Given the prominence of this theme in the interviews, we included English language proficiency as a variable of interest in the study. In addition, interviewees also proposed that membership in a freelancing union could challenge discrimination through a potential dispute resolution process where (a) the prices of labour were standardised to avoid undercutting, and (b) freelancers and clients could be reported to a union for discrimination and face sanctions on the platform. We thus included 'union membership' as one of our five design interventions. Table 3 outlines the design-based interventions that were selected for testing. We include a brief description of each intervention and the literature that the strategy was developed from. When a specific strategy was suggested by freelancers independently in interviews, this is noted in the source column with 'interview'.

Experimental study design and mock platform

The experimental study was conducted from September to October 2022 through the standalone web application hosted on *Mock-Freelancer.com* (see Figure 2). We recruited participants through Amazon Mechanical Turk (MTurk) which is often used to test new platforms. As MTurk workers participate in the gig economy, they are likely invested in our description of our study, to test a new freelancing platform. This allowed us to recruit participants quickly (Almaatouq et al., 2020). We assigned each user a unique confirmation code that allowed us to accomplish two key aspects of the experiment: pay the participants for completing their task and ensure that each participant was indeed unique. Participants in the experiment were provided instructions through a pop-up window which required confirmation of understanding before they could continue with the experiment. Participants could revisit instructions at any point during the experiment by clicking a 'Help' button. To align with the survey and interviews, participation in the experiment was narrowed to individuals based in the UK.

Following the procedure for a-priori analysis laid out by Stefanelli & Lukac (2020), we determined that a sample of 1500 hirers was adequate to uncover an effect of at least a 3-percent difference in choice propensity (*or preference*)

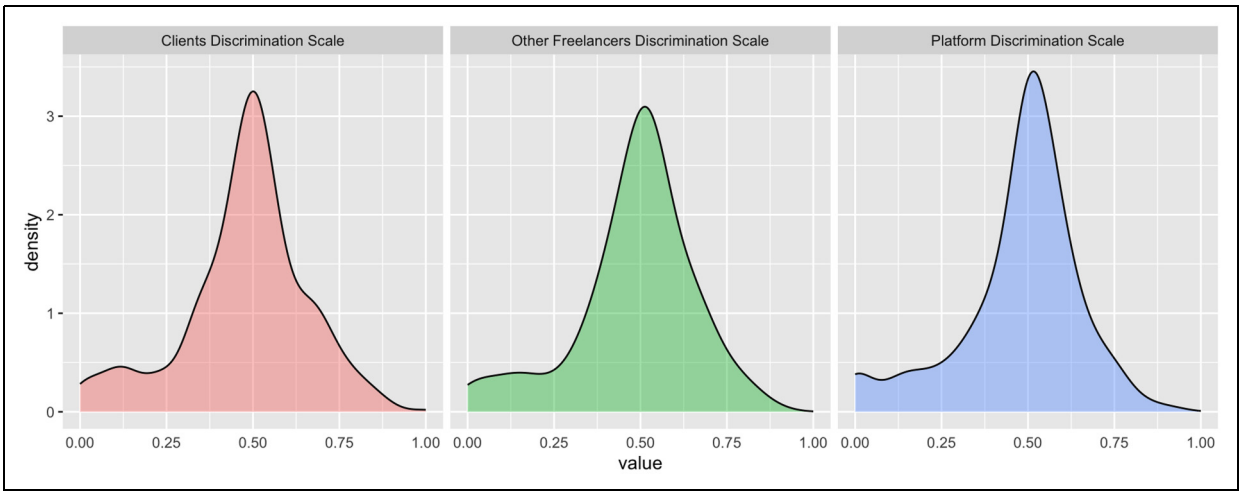

Figure 1. Density of discrimination from clients, colleagues, and freelancing platforms as reported in the survey.

Table 2. Summary interview schedule with example questions.

Interview theme	Example questions
Freelancing Background	How do you decide how much to charge for a project?
Perspectives on Discrimination	What do you think discrimination looks like in online freelancing? How common is discrimination?
Experiences of Discrimination	Do you think who a freelancer is (identity, race, gender) influences the prices they expect for their work? Do you think your identity affects the prices you can charge?
Role of Platforms	Do you think freelancing platforms encourage discrimination?
Interventions Feedback and Usability Testing	Interviewee interacts with the mock freelancing market. Provides feedback on the platform and interventions as they use it.
Ideas for Additional Features	Are there other interventions you would like us to test? Why?
Closing Remarks	Is there anything else you would like me to know?

Table 3. Description and source of discrimination intervention strategies.

Intervention	Description	Source
Community	Profile images of freelancers who successfully completed similarly projects are shown.	Leung, 2021; Levy & Barocas, 2017; Interviews
Flairs	Woman-owned business, black-owned business, etc. displayed under the profile.	Interviews, Levy & Barocas, 2017; Pedulla, 2014
Text	Text reviews replaces numeric ratings.	Garg & Johari, 2019; Rivera & Tilcsik, 2016; Wood et al., 2019
Union	Union membership is shown, and projects and freelancers can be reported to the Union.	Interviews, Rosenblatt, 2019; Wood & Lehdonvirta, 2021
Prompt	Pop-up stating discrimination is taken seriously. Known in web design as an interstitial.	Levy & Barocas, 2017

in the model. Discrete-choice analysis (DCA) is a stated preference method developed to understand and model choices between pairs of alternatives. In our experiment, individuals were presented with sets of experimentally designed alternatives (two freelancer profiles) where the attributes of interest that are likely to be salient for mechanisms of discrimination (outlined in Table 4) were manipulated. We considered discrimination as occurring when a hiring decision prioritised attributes not indicative of ability over those that are. Each hirer was presented with 12 choice tasks with two alternatives, allowing us to analyse a total of 18,160 decisions. Of the 1554 experiment participants, 97% provided demographic information; 51% of the sample identified as men and 49% identified as women. Additionally, 25 individuals listed that they prefer to self-describe their gender through free text. In

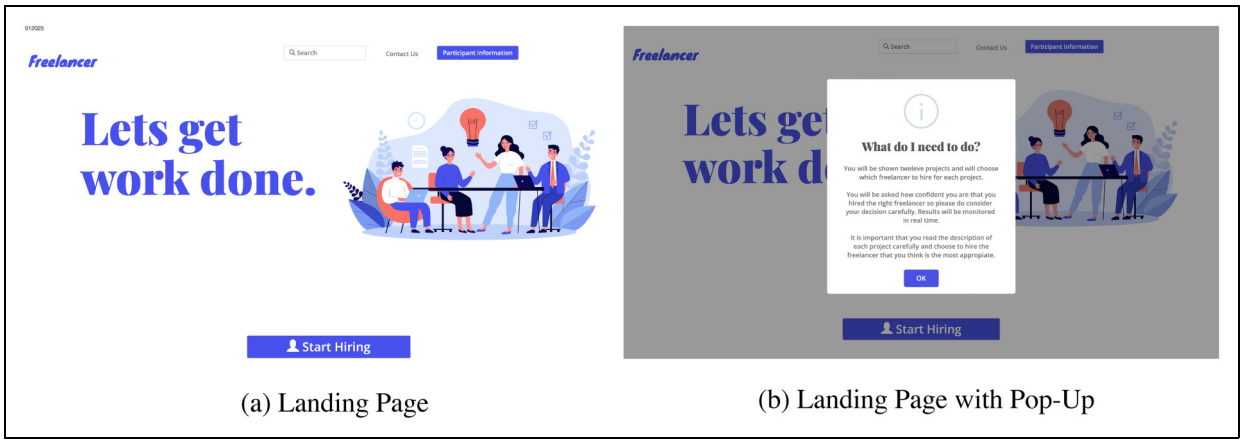

Figure 2. Mock freelancing platform landing page.

Table 4. Variables included in the research design.

	Variable	Levels and description
Experimental	Intervention	Equity interventions built into the Mock Freelancing Platform.
	Profile hired	*Binary [0, 1]*. Freelancer profile that the participant hires
Attribute Identity Features	Profile gender	*Categorical [Man, Woman]*. Gender of the freelancer. Inferred from profile image and display name.
	Profile ethnicity	*Categorical [White, Black, East Asian, South Asian]*. Ethnicity of the freelancer. Inferred from profile image.
	Proficiency	Native or Proficient English. Included from peer-researcher interviews.
Attribute Ability Features	Average rating	*Categorical [Low, Medium, High]*. Rating out of five for the freelancer.
	Number of reviews	*Categorical [Low, Medium, High]*. The number of reviews a freelancer has.
Participant Demographics	Participant gender	*Categorical [Man, Woman, Non-Binary, Gender-queer, Self-describe]*. The self-identified gender of the participant.
	Participant ethnicity	*Categorical [UK Survey Categories]*. The self-identified ethnicity of the participant.

terms of ethnicity, the sample identified as 92% White, compared to 82% of the UK. 4% of participants identified as Black, 3% as Asian (South and East), and 1% as belonging to Multiple Ethnic Groups. We followed the terminology of gender and ethnicity that is used in the UK Census data in 2021 throughout. The over-representation of White respondents led us to use a Chi Square Test of Independence to test if there was an association between hirers' ethnicity and the ethnicity of the freelancer they chose to hire. We compared the hiring decision of White respondents with non-White respondents. We found no significant difference based on hirer ethnicity. However, we remained sensitive towards an in-group bias towards hiring White freelancers (Almaatouq et al., 2020).

The jobs that formed each trial and intervention in *Mock-Freelancer.com* were informed by job descriptions posted on Upwork.com and other popular freelancing platforms. The experiment was conducted through a standalone web app developed in R Shiny.[2] We did not randomise the order of trials because of our concern that if the pop-up that directly addressed discrimination appeared first, it might prime hirers for the rest of their responses. Aligning the freelancers' listed skills with project specifications was designed to convey to hirers that both freelancers whose profiles they had to decide between had the skills required to complete the advertised project. Figure 2 shows the landing page of *Mock-Freelancer.com* and the instructional pop-up that appears when users click 'Start Hiring'. Figure 3 shows the general structure of the *Mock-Freelancer.com* choice pages for the baseline trials. Screen captures of the intervention strategies follow in Figure 4.

The freelancer profiles consisted of an image, user name, average rating, number of reviews, English language proficiency, location (all UK-based), and list of skills displayed as text on each freelancer's profile that mirrored the skills in the advert. Hirers were expected to infer the gender and ethnicity of freelancers from the images and usernames in freelancer profiles, a common practice in computational research on big data (Hofstra and De Schipper, 2018). The usernames were selected from the 100 most popular names in the UK census, followed by an underscore and up to four randomly generated numbers to simulate unique usernames. We used AI-generated images for

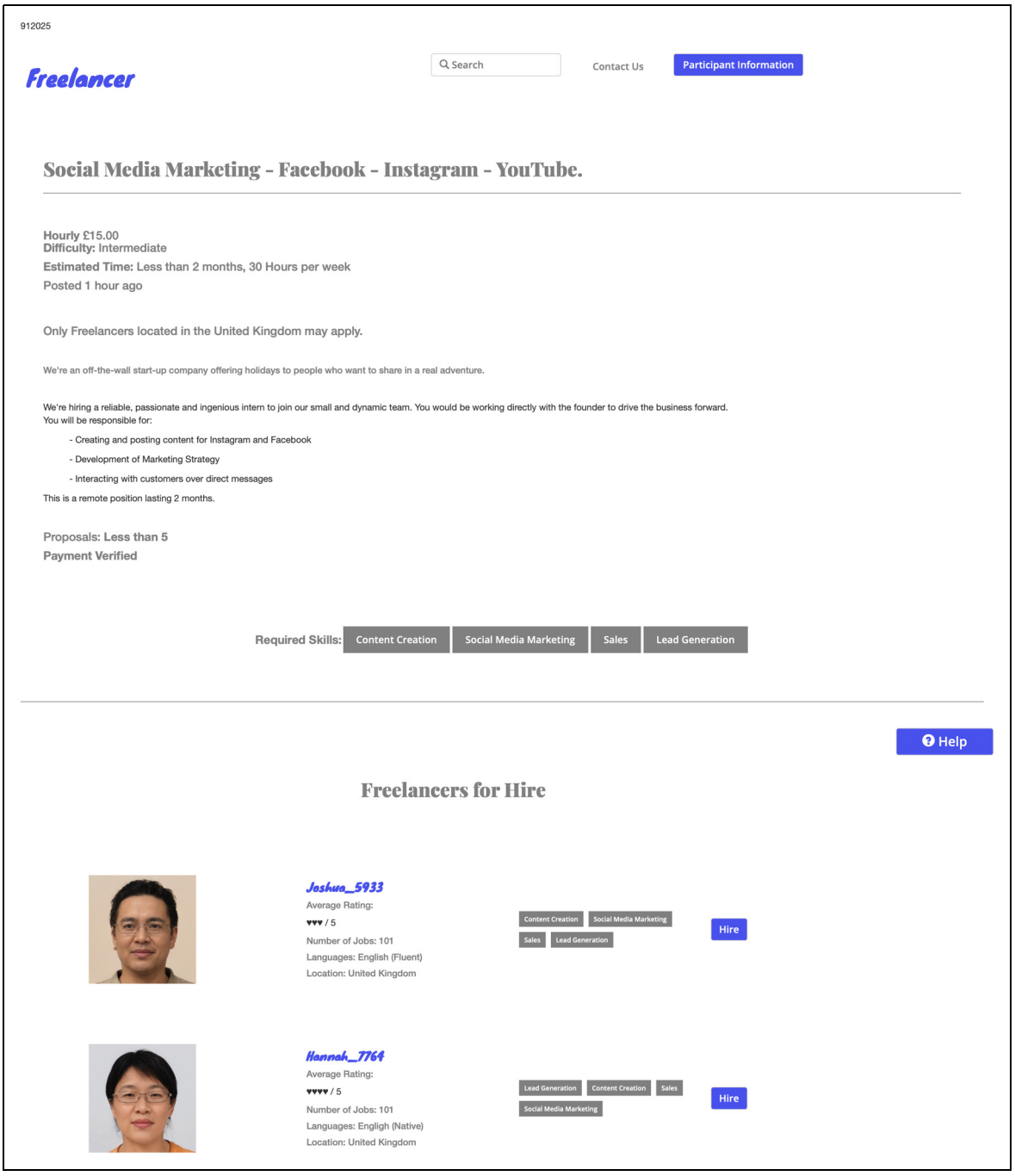

Figure 3. Example structure of the mock freelancing platform choice page.

freelancers' photos, selecting photos that appeared natural and were smiling to mimic typical freelancer profiles on OLMs.[3] We encoded the images to ensure no hirer saw the same freelancer more than once. For ratings, freelancers had low (1), medium (2.5), or high (5) scores that were depicted as stars out of a maximum of five stars. Accurately representing frequency as a categorical variable for the number of reviews was trickier. Instead, we populated the ratings with values that were within one standard deviation of the values 10 (Low), 50 (Medium), and 100 (High). Operationalising rating and reviews as categorical variables allowed us to compare freelancers, while retaining ecological validity by displaying numeric values to participants. The frequency of each category of ratings and

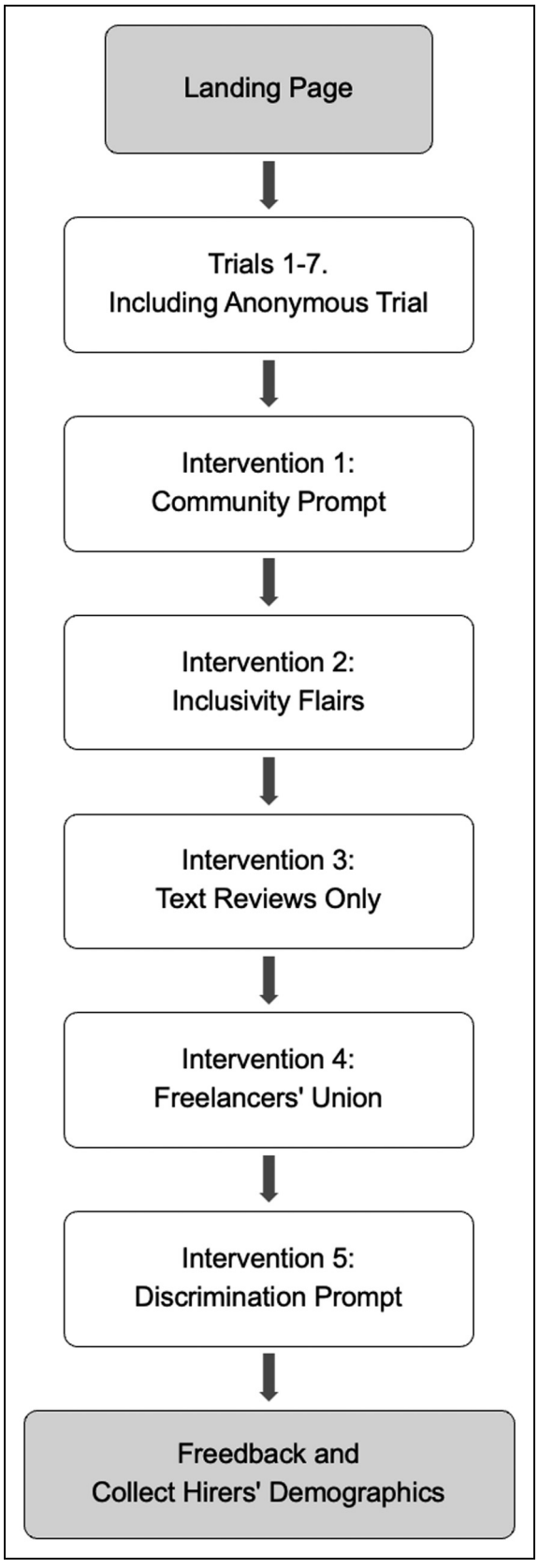

Figure 4. Structure and flow of the experimental research design.

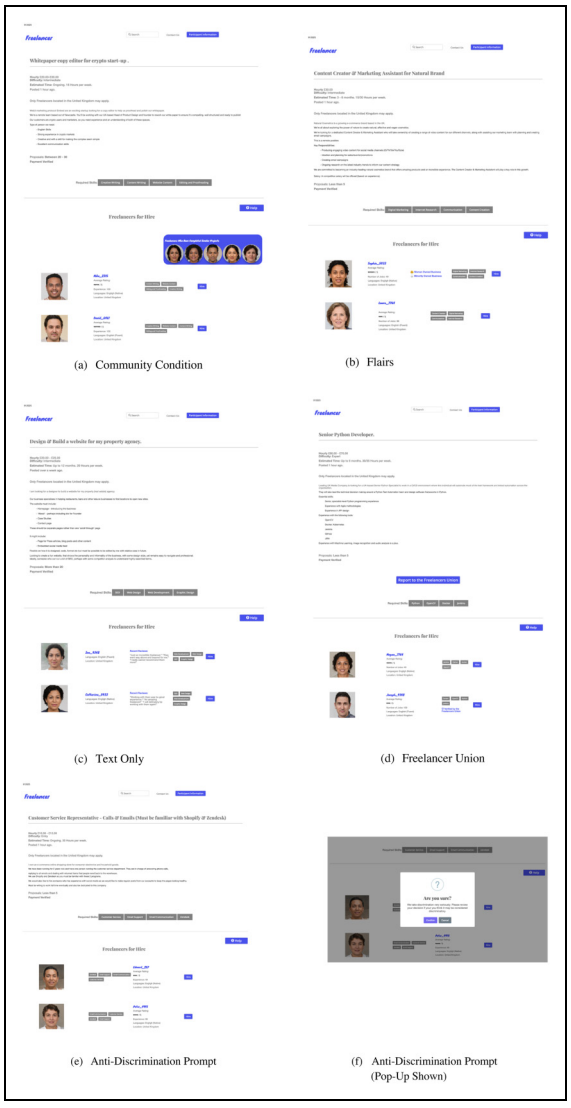

Figure 5. Screenshots of example intervention strategies.

reviews was matched proportionally across baseline and trials. The order in which freelancers were presented to hirers was not significant in post-hoc testing.

In addition to the attributes outlined in Table 4, we included an 'anonymous' freelancer profile. Our survey showed that freelancers may operate anonymously on freelancing platforms to avoid experiencing discrimination. This was coded as a dummy variable (0; Known attributes, 1; Anonymous), where a random known freelancer was compared with an anonymous profile. Both freelancers had identical skill sets that matched the job specifications displayed. The attributes in Table 4 reflect the components that make up each freelancer profile. Hierarchical Bayes (HB) was used to estimate the model parameters due to its ability to estimate individual-level parameters. Here, the model parameters are the part-worth utilities (shortened to 'utilities'), interpreted as preferences towards specific attribute levels,

Our implementation of DCA contained four sections: (a) Reference trials 1—7, a hiring choice without intervention for a baseline, (b) Interventions 1—5 shown in Table 3, (c) Feedback on the website design in a free-text box, and (d) Socio-demographic questions. The implementation of the interventions is shown in Figure 4. The experiment presented participants with seven discrete-choice tasks without interventions in the reference trials. The reference trials were structured to replicate the choices generated for the interventions to allow for comparison. The flow of the experimental research design is shown in Figure 5.

In the research design, we prioritised experimental orthogonality, ensuring that all specified parameters may be estimated independently of any other. The final model's variance inflation factor (VIF) scores indicated that the attributes were close to orthogonal, except for the Anonymous Dummy. A factional (or partial) research design was employed, using a representative subset of the complete set of possible combinations of attribute levels resulting in 144 freelancer profiles (of a potential 270), combined into choice sets of two as specified above. We used the polycor package to estimate the correlations between the factors and the VIF in Table 5.

Results

Freelancers experiences of discrimination and payment

Addressing RQ1, the survey examined if payment and experiences with freelancing platforms vary by gender and ethnicity. The mean day rate for freelancers in the survey was £154.20. There was no significant gender or ethnicity difference in the day rate for the freelancers who responded to the survey ($\alpha = 0.77$). Given that there is no 'reference' ethnicity, we used the Tukey Honest Significant Difference (HSD) test to compare the means of each ethnicity group and found no pairwise significant difference. While this is counter to previous research on offline labour markets, it is not a unique finding for OLMs (Foong and Gerber, 2021). Figure 6 shows the distribution of day rate by gender and ethnicity.

Our survey also looked at discrimination based on payment that might not be directly represented in the day rate of freelancers. We asked freelancers if they ever (a) had a discount requested, (b) had a discount of more than 50% requested, or (c) been asked to work on a project for free. This was broken down by ethnicity and gender through analysis of covariance (ANCOVA). The results of the ANCOVA for gender and ethnicity for requesting discounted labour are displayed in Table 6.

Table 5. Multicollinearity diagnostics (variation inflation factor).

Freelancer attribute	VIF
Gender	1.01
Ethnicity	1.03
English Language Proficiency	1.01
Average Rating	1.03
Number of Reviews	1.02

VIF: variance inflation factor.

Table 6 shows that while women are equally likely as men to have a discount requested, they are significantly more likely than men to be asked to do the work for free for 'experience'. Freelancers who are White or Black were significantly less likely to be asked for a discount. Post-hoc testing with Tukey HSD revealed significant differences in having any discount requested between White and Multiple Ethnicity freelancers ($t = -0.31$, $alpha = 0.00$**). No significant differences between White and Multiple Ethnicity freelancers ($t = -0.31$, $alpha = 0.00$**) and White and South Asian freelancers ($t = -0.41$, $alpha = 0.01$*) were found, where White freelancers had a discount of at least 50% requested less often. For being asked to work for free, there was a significant difference between White and Multiple Ethnicity freelancers ($t = -0.17$, $alpha = 0.00$***) and White and South Asian freelancers ($t = -0.32$, $alpha = 0.00$***) with White freelancers less likely to be asked to work for free. Overall, there are some correlations and significant differences in experiences of discounts being requested by gender and ethnicity. While there is no clear, overarching pattern, White freelancers appear to face fewer requests for reduced rates. Since payment rates are usually set by freelancers themselves, we decided to focus on the decision to hire, or not, in the experimental portion of this study.

Figure 7 shows the frequency of specific experiences of discrimination. Only 20% of freelancers reported never having any of the negative experiences detailed in the survey; 80% of freelancers surveyed had reported at least one instance of discrimination in the last year. Of those who had negative experiences, 76% of survey respondents attributed these experiences to their status characteristics. Our survey indicated that most freelancers turned to the relevant freelancing platform for assistance when they experienced discrimination.

Baseline aggregated preferences towards hiring freelancers

We sought to categorise intersectional hiring discrimination in OLMs, adding depth to RQ1. Using the HB method, we assessed the utility parameters for each hirer in the sample, as well as the entire sample. Table 7 shows the average part-worth utilities for each level of each attribute and the importance scores for each of the attributes. Part-worths explain the contribution of the levels of each attribute to the overall utility, allowing one to understand how

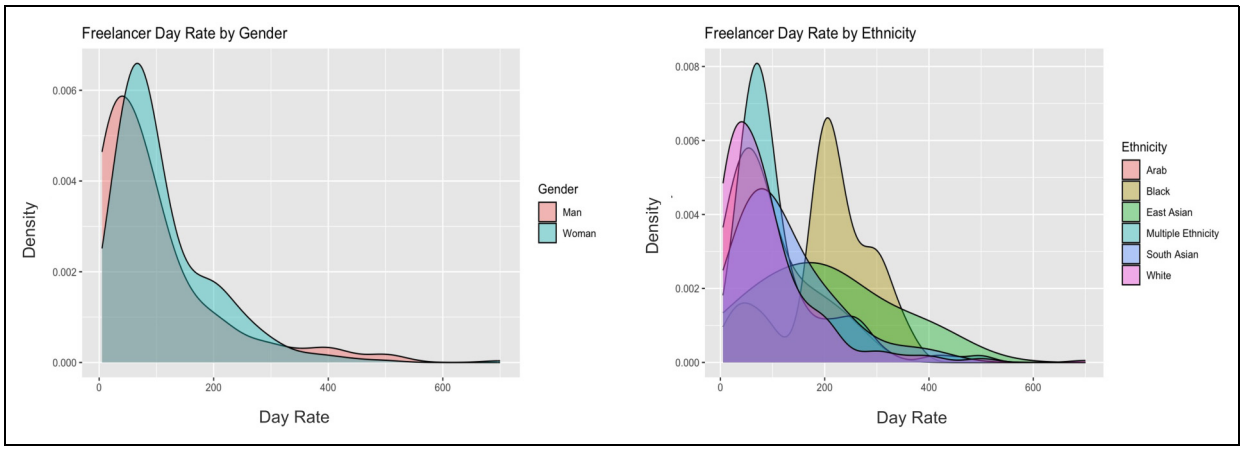

Figure 6. Freelancer day rate by gender and ethnicity from survey.

Table 6. Frequency of requested discount by gender and ethnicity.

	Requested discount	Requested >50% discount	Free
Gender: Woman (Reference: Man)	−0.04	0.01	0.08**
	(0.09)	(0.09)	(0.04)
Ethnicity: Black	−0.69**	0.07	0.10
	(0.35)	(0.35)	(0.17)
Ethnicity: East Asian	−0.62	0.18	−0.00
	(0.40)	(0.40)	(0.19)
Ethnicity: Multiple Groups	−0.23	0.37	0.06
	(0.30)	(0.30)	(0.15)
Ethnicity: South Asian	−0.37	0.47	0.21
	(0.31)	(0.31)	(0.15)
Ethnicity: White	−0.60*	0.07	−0.08
	(0.30)	(0.30)	(0.15)
Constant	3.04***	2.36***	1.63***
	(0.30)	(0.30)	(0.15)
Observations	667	666	652
F Statistic	3.02*** (df = 6; 66)	3.30*** (df = 6; 659)	7.16*** (df = 6; 645)
*$p < 0.1$; **$p < 0.05$; ***$p < 0.01$			

changing specific characteristics of the freelancer affects hirers' preferences for a particular hiring choice. Higher part-worth utility values indicate a greater preference for hiring a freelancer with that particular attribute. The part-worth utilities of each attribute are scaled to sum to 1, so they should be considered as relative values within the experimental setting. Figure 8 includes a non-scaled plot of the part-worths of each attribute.

Table 7 shows that the average rating had the most significant influence on hirers' preferences, with a relative importance of 41%. The number of reviews was the second most important attribute, responsible for 32% of the hiring decision. As expected, hirers preferred freelancers with the highest average rating and number of reviews. Notably, there is a greater difference between a medium and high number of reviews than between a medium and low number of reviews. There is thus a relative advantage for a high number of reviews compared to a medium than a medium number of reviews compared to low. Freelancer ethnicity (as a combined attribute) was reasonably salient, contributing 14% of the choice to hire. The preferred freelancer is South Asian. Hirers showed the least preference for freelancers who are East Asian. We find a preference for native English speakers over fluent speakers, which seems reasonable given the experiences freelancers cited in our interviews. Hirers in this study generally preferred women, and freelancer gender explained 5% of the overall hiring decision, as shown in Table 7. This supports previous research (i.e. Chan and Wang, 2018), which also suggest that in contrast to hiring in offline labour markets, hiring on OLMs favours women.

We excluded comparisons with anonymous freelancers who had no discernible gender or ethnicity visible on their profile. The survey found that 76% of freelancers

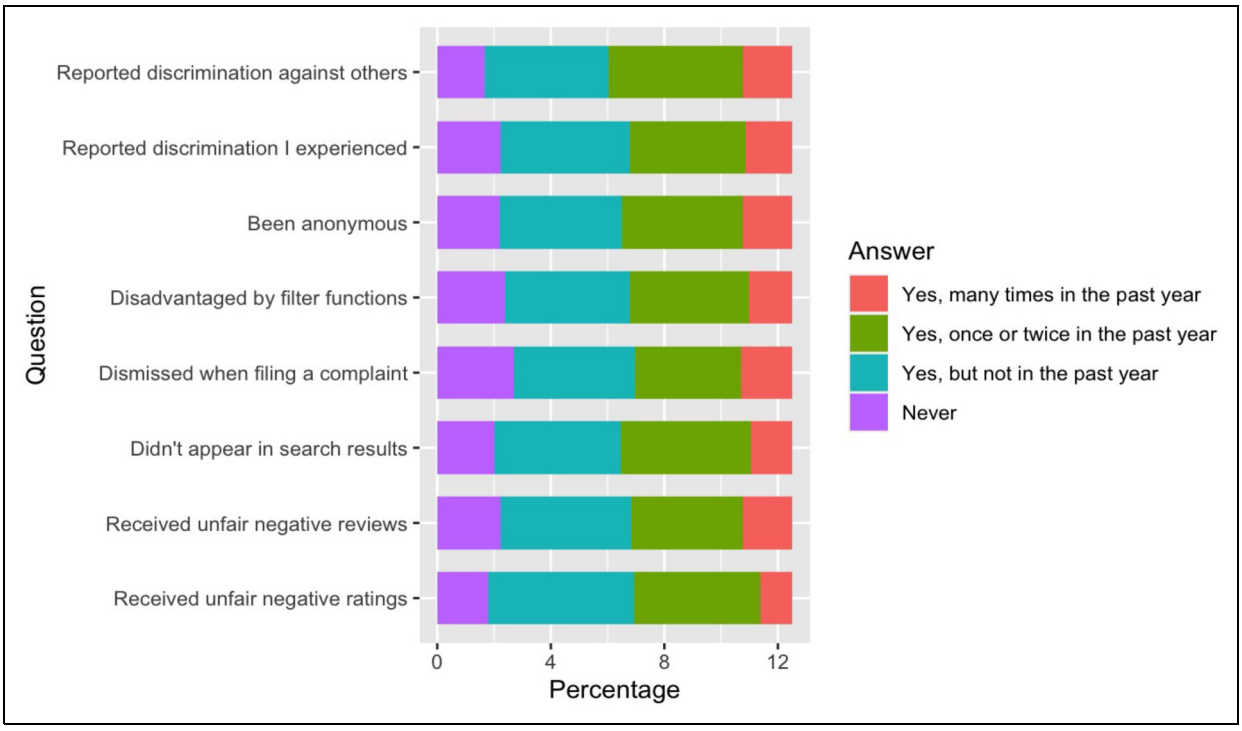

Figure 7. Distribution of specific experiences of platform discrimination from survey.

Table 7. Importance of each freelancer attribute in the decision to hire (baseline attribute part-worth).

Attribute	Baseline part-worth
Gender	0.05
Ethnicity	0.14
English Language Proficiency	0.10
Average Rating	0.41
Number of Reviews	0.32
Base utility	0.59

conduct business under their given name, with 51% using an image of themselves and 48% also using a logo. The figure below indicates how anonymity compared to the other possible combinations of gender and ethnicity. Anonymity generally disadvantages freelancers. Figure 9 shows that compared to being anonymous, a freelancer who shows their gender is preferred by hirers. Anonymity only benefits freelancers who are East Asian men. This aligns with several interviewees reflecting that they may be hired from an anonymous profile but be removed from a project when they appeared in a video call.

Finally, we examined the selected platform-based interventions (Table 3) that affect hiring decisions and discrimination in OLMs (RQ2). Table 8 demonstrates a comparison of the utility of each attribute under each intervention with the baseline model. The baseline and intervention models all passed significant testing measures. Across all interventions, hirers preferred a high number of reviews and average rating. A-priori power analysis indicated sufficient hirers under each condition.

Under the community condition, English language proficiency (20%) and ethnicity (17%) increase in importance for hirers' decision-making. While the average rating and number of reviews are relatively more important factors, their influence decreased by close to 8%. Under this intervention, there is a preference for Black freelancers, with East Asian, South Asian, and White freelancers negatively affected. In this intervention, hirers were shown a community of '*Freelancers who had successfully completed similar tasks'*, composed of non-White individuals. This indicates that community composition favours the hiring of freelancers who share identity attributes with freelancers who have already successfully completed a similar task.

Under the flairs condition (i.e. 'Woman Owned Business'), ethnicity (35%) and gender (34%) are the most important attributes, followed by the average rating (19%), number of reviews (9%), and English language proficiency (1%). For ethnicity, hirers strongly preferred South Asian (18%) followed by East Asian (11%) freelancers. For gender, the preference for women rose to 18%, a significant increase of 17% from the baseline model. Inclusivity Flairs increased the preference for hiring women and countered the bias against East Asian freelancers. Flairs present a potential intervention for design justice in OLMs.

In our text reviews condition, we examine how text-based reviews affect attribute preferences in hiring decisions compared to the baseline model. This intervention only allows a

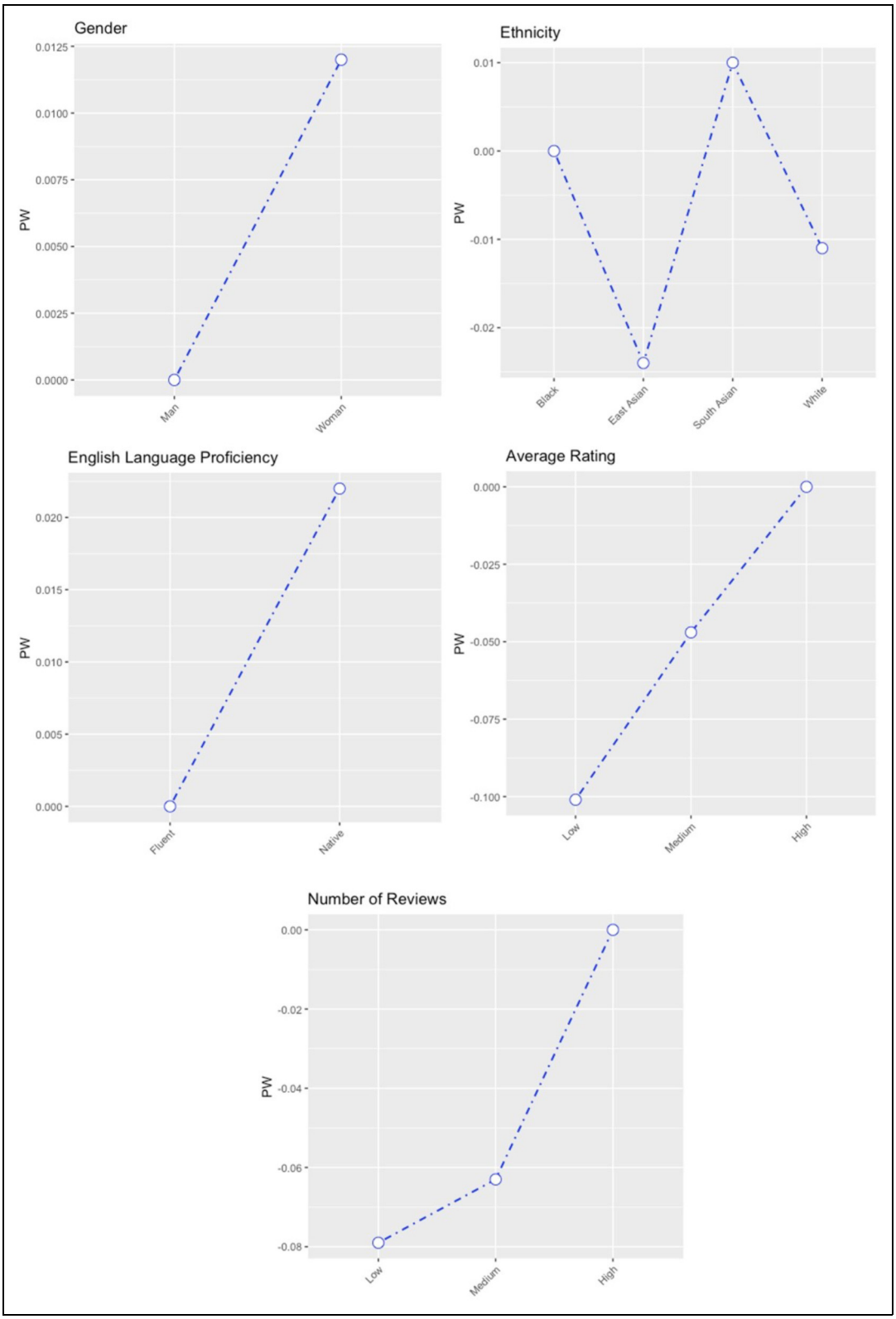

Figure 8. Baseline attribute part-worths in hiring mock freelancer.

limited comparison with the intervention model, as the two attributes – number of reviews and average rating – are replaced with equivalent text reviews. However, the model has interesting findings. Under the text reviews condition, ethnicity is the strongest indicator of preference, followed by gender and English language proficiency. Under the text

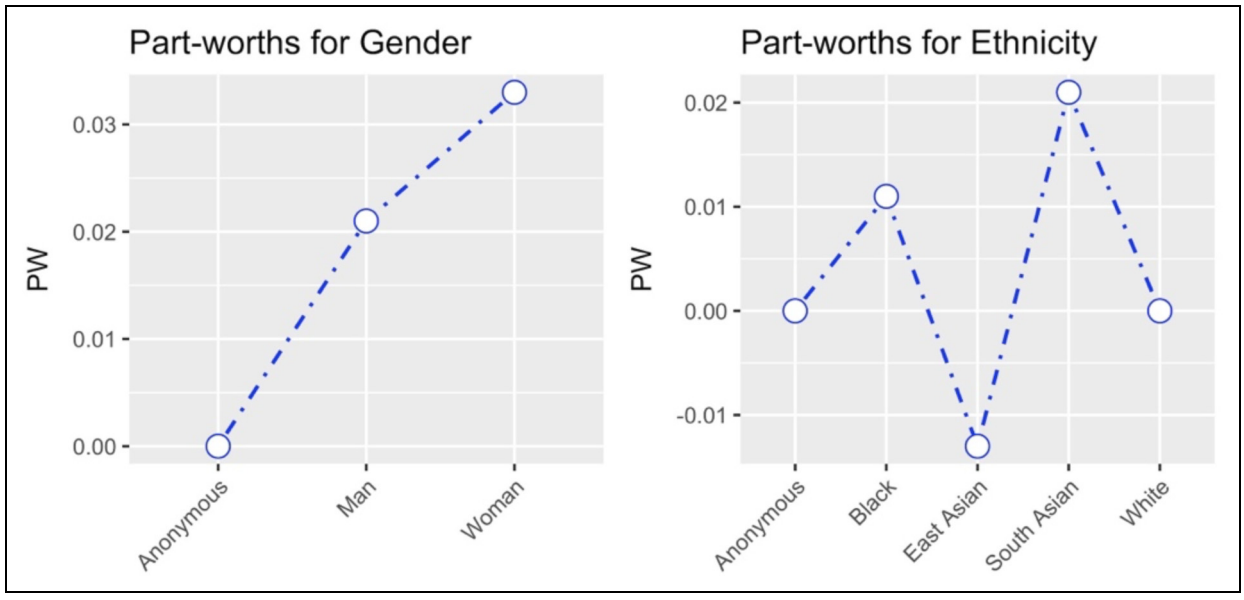

Figure 9. Baseline gender and ethnicity part-worths, including anonymity.

Table 8. Comparison of attribute part-worth under each intervention strategy.

Attribute	Baseline	Community	Flairs	Text	Union	Prompt
Gender	0.05	0.04	0.34	0.30	0.06	0.22
Ethnicity	0.14	0.18	0.35	0.42	0.17	0.27
English Language Proficiency	0.09	0.20	0.01	0.29	0.16	0.00
Average Rating	0.41	0.32	0.20	-	0.31	0.22
Number of Reviews	0.32	0.26	0.10	-	0.31	0.29
Base utility	0.59	0.53	0.35	0.44	0.53	0.51

reviews condition, hirers indicated a preference for South Asian freelancers, followed by Black freelancers. White freelancers were the least preferred. Women remained preferred. Without a quantitative basis for comparison, ethnicity appears to become the strongest indicator of freelancer hiring.

Under the 'union membership' intervention, where freelancers' profiles indicated that they were members of a Freelancers' Union, average rating, and number of reviews (31%) were equally important for hirer's preferences. Ethnicity was also important, responsible for 17% of the hirer's decision. There was a preference for East Asian freelancers. Notably, there was also an increased preference for White freelancers by 5%. Despite the overrepresentation of White hirers in our sample, an increase in preference is noteworthy because it demonstrates a shift from the reference model; 16% of the preference can be attributed to English Language Proficiency. The shift in preferences for freelancers could be due to an assumed decrease in 'undercutting' prices that freelancers in the interviews said they would benefit from, with the Union assisting with standardising rates. The Union intervention upends the bias experienced by East Asian freelancers, compared with the baseline hiring model.

Under the discrimination prompt, which asked users if they would like to review their hiring decision, the number of reviews remained most important (29%), followed by ethnicity (27%). There was an increased preference for South Asian freelancers, with a bias against East Asian and White freelancers. There remained a preference for women, with gender responsible for 22% of the choice. Actively challenging potentially discriminatory behaviour may increase the importance of ethnicity in hiring.

Design intervention techniques are most effective at increasing the importance of ethnicity and gender when used as tools for positive discrimination or *justice*, undoing allocative and representational harm (Davis, Williams and Yang, 2021). Such justice is evidenced in the community and flairs intervention. In addition, interventions which directly address discrimination, such as the discrimination prompt, should be used with caution since they appear to amplify ethnicity in hiring decisions.

Discussion and conclusions

Affordances are critical in shaping how technologies can be best marshalled to minimise discrimination that often

characterises social and economic life. Prior studies have identified how design features of platforms shape the affordances of those platforms (Schwartz and Neff, 2019; Stringhi, 2022), including in how platforms impact worker autonomy (Wood et al., 2019; Bellesia et al., 2013). We have extended this line of research to manipulate platform design features precisely to better align design justice principles to practise. The mechanisms of affordance that we focused on were designed to affect what Davis (2020: p.65) defines as the 'condition of cultural and institutional legitimacy' or how larger social-structure norms inform human-technology relations. By default, technologies tend to reproduce existing power structures. To subvert these structures, we collaborated with freelancers to develop design interventions (affordance *mechanisms*) that would challenge the structures of power that perpetuate discrimination in OLMs.

As our initial survey of freelancers illustrated, the norm for freelancers is to operate on OLMs under their personal identity. The rating and review systems of OLMs are typically assumed to be a neutral mechanism conveying a freelancer's ability and quality of previous work to a potential hirer. Indeed, our experiment found that a freelancer rating and number of reviews were responsible for the majority (72%) of the decision to hire in the baseline. Prior studies have noted the importance of reputation in OLMs, as ascertained by numerical ratings and especially by text reviews (Pallais, 2014). In the absence of other information, reputation appears to stand in as a testament to the skills and reliability of a freelancer. With the inclusion of text reviews as a mechanism for intervention, White freelancers were least preferred. Hannak et al. (2017) found that freelancers who were not White generally received less text reviews on OLMs. We find that that an equal volume of text reviews between White and non-White freelancers benefits the latter. In the context of our experiment, it is possible that ethnicity may simply be far more visible in the choices that hirers make. Hirers may thus be correcting for racial biases they see themselves as having, with text reviews providing the appropriate mechanism of affordance where hirers feel they can confidently ascertain a freelancer's skill and suitability, enabling hirers to make bias-correcting decisions in a less risky context. Given rating inflation in OLMs, it is likely that the number of reviews and ethnicity of freelancer will influence hiring preferences. Hannák et al. (2017) find that perceived ethnicity is significantly correlated with ratings and text evaluations. In one U.S.-based study of TaskRabbit and Fiverr, researchers found that Black and (East) Asian workers received lower ratings and more negative adjectives. However, the ethnic group that was ranked lowest actually shifted based on geographic location within the US (Hannák et al., 2017). Hirers do not treat 'non-White' groups as an undistinguishable monolith, but have their own hierarchies. In our UK-based study, our findings suggest that South Asian freelancers benefit from rating inflation, and possibly at a cost to East Asian freelancers, since even text reviews do not seem to benefit East Asian freelancers.

Scholars have suggested that being anonymous on OLMs could mitigate race and gender biases, for instance, in terms of preferring women for lower-value projects or minimising the disadvantage for East Asian freelancers. Unlike offline labour markets, it is ostensibly easier for freelancers to be anonymous on platforms if they choose, for instance, to forgo a profile picture or share a handle in place of their name. We found that anonymity was generally undesirable for freelancers, only increasing the likelihood of being hired for East Asian freelancers, who otherwise tend to be least preferred for hiring. This somewhat counter-intuitive finding may be because of a norm in OLMs to share considerable information to humanise oneself on the platforms. Anonymous profiles appear to go against the cultural and institutional condition of affordances (Davis, 2020), with OLMs functioning as spaces that expect some level of information about freelancers. This may come down to trustworthiness, which hirers may be searching for as they quickly examine freelancer profiles. It could also be an extension of the expectation in offline labour markets to establish a sense of 'chemistry' with employers by emphasising one's personal attributes, and the cultural expectations of bringing passion to one's work (Rao and Tobias Neely, 2019; Sharone, 2013). The contemporary cultural context of work appears to discourage anonymity, even on OLMs. That said, some platforms operate on algorithmic matching systems, such as Uber, where drivers appear indistinct as do riders. Why anonymity works for some in OLMs, but not others is an avenue of further inquiry that may be fruitfully pursued through future scholarship.

Flairs, such as a decorative logo denoting a business as 'minority-owned', are also crucial in minimising discrimination in OLMs. Flairs lead hirers in our study to choose women and non-White freelancers. Given the disadvantage that East Asian freelancers (especially East Asian men) encounter, it is possible that an appropriately designed flair could mitigate these biases within OLMs. Our data are not suitable for explaining the mechanisms through which flairs counter prejudices against East Asians. One mechanism could be that in the context of similar choices, flairs serve as a low-cost but feel-good decision wherein hirers may feel that they are behaving beneficently by selecting the freelancer in question.

Our study has explored how interventions for design justice can impact hiring decisions. In offline studies, discrimination against women is especially rife in elite labour markets where decision-makers' biases about women's suitability for particular occupations are

foregrounded (Rivera and Tilcsik, 2016). Freelancers are often better educated and come from a more privileged background than the average population. Yet the jobs they are typically vying for through OLMs are not elite jobs with expectations of stability, benefits, and even some semblance of security. The projects offered on OLMs, even for highly educated workers, are usually time-limited. The characteristics of the job in question may thus explain the differences in mechanisms of discrimination. The duration of many of these jobs and other such aspects may lead hirers to privilege women, especially Asian women, who are often stereotyped as detail-oriented, diligent, and good at time and organisational management (Leung, 2018).

Based on our findings, we offer the following recommendations for platform designers and researchers:

Include Freelancers: It is imperative that platforms include freelancers and their experiences in developing features to intervene in discrimination. Through our interviews with freelancers, we learned about intervention mechanisms to potentially minimize discrimination that were not prioritised in the literature, such as union membership and English language proficiency. Understanding freelancers' experiences of discrimination provides vital insight into biases that matching algorithms can potentially learn and exacerbate.

Flairs and Reviews: Platforms can deploy justice interventions that raise the visibility of users who typically face discrimination. In our study, these interventions are community composition and inclusively flairs. OLMs could remedy this by allowing freelancers to choose up to a predetermined number of reviews (i.e. five) for their profile, selected from all their reviews. This allows freelancers some negotiation in how there are perceived on the platform.

Community Nudges: Show hirers that a diverse spectrum of freelancers have successfully completed similar tasks. Our findings show that community interventions can challenge gender and ethnicity preferences, which may affect the chances of a freelancer being hired.

Caution with Explicit Interventions: Specific prompts on discrimination should be used with caution as they appear to encourage the importance of ethnicity in decision making.

Guided by the conceptual framework of affordances and scholarship on design justice, our methodological contribution is creating a mock OLM that allows us to experimentally test platform design interventions. Such experimentation has not been possible in prior studies because typically platform owners, for example Upwork, Freelancer, TaskRabbit and so on, tend to protect details about the inner functioning of their platforms. Scholars have repeatedly noted the lack of transparency of how platform algorithms are produced (Baumer, 2017; Jarrahi et al., 2021; Noble, 2018). Our research design allows us to innovatively overcome this challenge. However, there are limitations to this study. Most importantly, in an experimental procedure the results occur in a tightly controlled research setting. We are thus cautious about any claims to generalisability. For example, our study modelled the choice between hiring one of two freelancers and limiting the information available to hirers. While choosing between two freelancers does not represent OLMs such as Upwork.com, it allows us to model preferential decisions in a hiring context. Leung (2021) highlights how competition in freelancing can cause a 'race to the bottom' where a platform like Upwork can downgrade the rate of pay. There appears to be an assumption that South and East Asian workers will inevitably accept fees lower than their White counterparts for the same work, coupled with a disadvantage based on assumptions of being non-native English speakers (Leung, 2021). Our research provides statistical evidence to support these assertions. Previous scholarship also suggests that stereotypes influence hiring decisions in freelancing, with hirers preferring workers whom they can pay a lower rate to complete the task (Galperin, 2021). As this study did not include price, our findings cannot account for these theories explicitly.

Future research should model if interventions can affect the price a freelancer is awarded for a project. While our data do not allow us to infer the mechanisms for this preference, other studies have suggested that gendered assumptions may lead to these observed outcomes (Cerqua and Urwin, 2018; Galperin, 2021; Hong, 2016). Our study could be expanded upon in the future by testing if South and East Asian freelancers may be preferentially hired due to assumed lower rates. Given the significant effects of community, flairs, and text interventions, our study indicates that more research on user-interface is warranted. For instance, researchers should examine how hirers interact with filtering mechanisms on OLMs to find freelancers and if identity-signalling flairs could assist with designing justice on OLMs. Future research should complement investigations of algorithmic structures with a critical analysis of how platform design communicates affordance to its users. Our study can additionally be expanded upon with a natural experiment, such as that conducted by Hangartner et al. (2021), to ascertain if the findings are repeated in a natural setting. How gender and ethnicity biases are reflected in these metrics and in freelancers' payment for their work will be a fruitful avenue for future research. This study provides experimental evidence of paths forward that can enable OLMs to function in a more equitable manner.

Acknowledgements

The authors thank the interview participants for their willingness to speak with us and for providing critical insights. The authors would additionally like to recognise the editor and anonymous reviewers for their contributions to this work. We would also like to acknowledge Fabian Stephany, Julian Albert, Vili Lehdonvirta, Martin Lukac for their valuable feedback to this study. We thank Nayana Prakash for her work as Research Assistant on this project.

Declaration of conflicting interests

The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This work was supported by the British Academy/Leverhulme Small Research Grant (SRG21\210549) and The Suntory and Toyota International Centres for Economics and Related Disciplines.

ORCID iDs

Siân Brooke https://orcid.org/0000-0001-7558-7924
Aliya Hamid Rao https://orcid.org/0000-0003-0674-4206

Notes

1. https://www.statista.com/statistics/318234/united-kingdom-self-employed/
2. A simplified version of the mock freelancing platform can be accessed here: https://sjmbrooke.shinyapps.io/discOLM/. Please note that the hosted platform has been amended for ongoing memory efficiency.
3. https://generated.photos/faces

References

Almaatouq A, Krafft P, Dunham Y, et al. (2020) Turkers of the world unite: Multilevel in-group bias among crowdworkers on Amazon mechanical Turk. *Social Psychological and Personality Science* 11(2): 151–159.

Aneja A, Luca M and Reshef O (2023) *Black Ownership Matters: Does Revealing Race Increase Demand For Minority-Owned Businesses? (No. w30932)*. Cambridge, Massachusetts: National Bureau of Economic Research. http://www.nber.org/papers/w30932.

Babar Y, Mahdavi Adeli A and Burtch G (2023) The effects of online social identity signals on retailer demand. *Management Science* 69(12): 7335–7346.

Baumer EP (2017) Toward human-centered algorithm design. *Big Data & Society* 4(2): 205395171771885.

Bellesia F, Mattarelli E and Bertolotti F (2023) Algorithms and their affordances: How crowdworkers manage algorithmic scores in online labour markets. *Journal of Management Studies* 60(1): 1–37.

Big Freelancer Survey (September 2022) Open To All, But Not All Hours. Future of Creative Work group in the Centre for Work, Organization and Society at the University of Essex. https://freelancersmaketheatrework.com/wp-content/uploads/2022/09/Big-Freelancer-Survey-2-report-FINAL.pdf.

Brooke SJ (2021) Trouble in programmer's paradise: Gender-biases in sharing and recognising technical knowledge on stack overflow. *Information, Communication & Society* 24(14): 2091–2112.

Cerqua A and Urwin PJ (2018) Unpicking the gender hiring bias in online labor markets. *Academy of Management Proceedings* 2018(1): 16137.

Chan J and Wang J (2018) Hiring preferences in online labor markets: Evidence of a female hiring bias. *Management Science* 64(7): 2973–2994.

Correll SJ (2017) SWS 2016 Feminist lecture: Reducing gender biases in modern workplaces: A small wins approach to organizational change. *Gender & Society* 31(6): 725–750.

Costanza-Chock S (2020) *Design Justice: Community-Led Practices to Build Worlds We Need*. Cambridge, MA: MIT Press. ISBN: 9780262043458.

Davis JL (2020) *How Artifacts Afford: The Power and Politics of Everyday Things*. Cambridge, United States: MIT Press.

Davis JL, Williams A and Yang MW (2021) Algorithmic reparation. *Big Data & Society* 8(2): 205395172110448.

Dekker R, Koot P, Ilker Birbil S, et al. (2022) Co-designing algorithms for governance: Ensuring responsible and accountable algorithmic management of refugee camp supplies. *Big Data & Society* 9(1): 205395172210878.

Eubanks V (2018) *Automating Inequality: How High-Tech Tools Profile, Police, And Punish The Poor*. New York, USA: Macmillan. ISBN: 9781250074317.

Foong E and Gerber E (2021) Understanding gender differences in pricing strategies in online labor marketplaces. In: *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*. Yokohama, Japan: ACM, 1–16.

Gaddis SM (2015) Discrimination in the credential society: An audit study of race and college selectivity in the labor market. *Social Forces* 93(4): 1451–1479.

Galperin H (2021) "This Gig Is Not for Women": Gender stereotyping in online hiring. *Social Science Computer Review* 39(6): 1089–1107.

Garg N and Johari R (2019) Designing informative rating systems: Evidence from an online labor market. *Manufacturing & Service Operations Management* 23(3): 589–605.

Goldin C and Rouse C (2000) Orchestrating impartiality: The impact of 'Blind' auditions on female musicians. *The American Economic Review* 90(4): 715–741.

Hangartner D, Kopp D and Siegenthaler M (2021) Monitoring hiring discrimination through online recruitment platforms. *Nature* 589(7843): 572–576.

Hannák A, Wagner C, Garcia D, et al. (2017) Bias in online freelance marketplaces: Evidence from TaskRabbit and Fiverr. In: *Proceedings of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing*. Portland, Oregon, USA: ACM, 1914–1933.

Hofstra B and De Schipper NC (2018) Predicting ethnicity with first names in online social Media networks. *Big Data & Society* 5(1): 205395171876114.

Hong R (2016) Soft skills and hard numbers: Gender discourse in human resources. *Big Data & Society* 3(2): 205395171 667423.

International Labour Organisation (2021) *Work Employment and Social Outlook: The role of digital labour platforms in transforming the world of work*. ILO Flagship Report. Geneva, Online Labour Observatory. https://www.ilo.org/global/research/global-reports/weso/2021/WCMS_771749/lang–en/index.html.ISBN: 9789220319444

Jarrahi MH, Newlands G, Lee MK, et al. (2021) Algorithmic management in a work context. *Big Data & Society* 8(2): 205395 172110203.

Kang M (2010) *The Managed Hand: Race, Gender, and the Body in Beauty Service Work*. University of California Press.

Kässi O and Lehdonvirta V (2018) Online labour index: Measuring the online gig economy for policy and research. *Technological Forecasting and Social Change* 137(1): 241–248.

Lehdonvirta V, Oksanen A, Räsänen P, et al. (2021) Social media, web, and panel surveys: Using non-probability samples in social and policy research. *Policy & Internet* 13(1): 134–155.

Leung WF (2018) *Digital entrepreneurship, gender and intersectionality: An east Asian perspective*. Switzerland: Springer.

Leung WF (2021) Freelancing globally: Upworkers in China and India, neo-liberalisation and the new international putting-out system of labour (NIPL). In: *Work and Labour Relations in Global Platform Capitalism*. Cheltenham, UK: Edward Elgar Publishing, 134–156. https://doi.org/10.4337/9781802205138.00015.

Levy K and Barocas S (2017) Designing against discrimination in online markets. *Berkeley Technology Law Journal* 32(3): 1183–1238.

Noble SU (2018) *Algorithms of Oppression: How Search Engines Reinforce Racism*. New York: New York University Press.

Pallais A (2014) Inefficient hiring in entry-level labor markets. *American Economic Review* 104(11): 3565–3599.

Pedulla DS (2014) The positive consequences of negative stereotypes: Race, sexual orientation, and the job application process. *Social Psychology Quarterly* 77(1): 75–94.

Quadlin N (2018) The mark of a woman's record: Gender and academic performance in hiring. *American Sociological Review* 83(2): 331–360.

Rao AH and Neely MT (2019) What's love got to do with it? Passion and inequality in white-collar work. *Sociology Compass* 13(12): e12744.

Ridgeway CL and Cornell SJ (2006) Consensus and the creation of status beliefs. *Social Forces* 85(1): 431–453.

Rivera L and Tilcsik A (2016) Class advantage, commitment penalty: The gendered effect of social class signals in an elite labor market. *American Sociological Review* 81(6): 1097–1131.

Rivera LA and Tilcsik A (2019) Scaling down inequality: Rating scales, gender bias, and the architecture of evaluation. *American Sociological Review* 84(2): 248–274.

Rosenblatt A (2019) *Uberland: How Algorithms Are Rewriting the Rules of Work*. Oakland, CA: University of California Press. ISBN: 9780520324800.

Schwartz B and Neff G (2019) The gendered affordances of craigslist "New-in-Town Girls Wanted" ads. *New Media & Society* 21(11–12): 2404–2421.

Sharone O (2013) *Flawed System/Flawed Self: Job Searching and Unemployment Experiences*. Chicago: University of Chicago Press.

Stefanelli A and Lukac M (2020) Subjects, trials, and levels: Statistical power in conjoint Experiments. *SocArXiv*: 1–24. DOI: 10.31235/osf.io/spkcy.

Stephany F, Kässi O, Rani U, et al. (2021) Online labour index 2020: New ways to measure the world's remote freelancing market. *Big Data & Society* 8(2): 2053951721 10432.

Stringhi E (2022) Addressing gendered affordances of the platform economy: The case of UpWork. *Internet Policy Review* 11(1). DOI: 10.14763/2022.1.1634.

Tak E, Correll SJ and Soule SA (2019) Gender inequality in product markets: When and how status beliefs transfer to products. *Social Forces* 98(2): 548–577.

Viljoen S, Goldenfein J and McGuigan L (2021) Design choices: Mechanism design and platform capitalism. *Big Data & Society* 8(2): 205395172110343.

Wagner G, Prester J and Paré G (2021) Exploring the boundaries and processes of digital platforms for knowledge work: A review of information systems research. *The Journal of Strategic Information Systems* 30(4): 101694.

Wood AJ, Graham M, Lehdonvirta V, et al. (2019) Good gig, bad gig: Autonomy and algorithmic control in the global gig economy. *Work, Employment and Society* 33(1): 56–75.

Wood AJ and Lehdonvirta V (2021) Antagonism beyond employment: How the "Subordinated Agency" of labour platforms generates conflict in the remote gig economy. *Socio-Economic Review* 19(4): 1369–1396.