Exploring people’s perceptions and support of data-driven technology in times of COVID-19: the role of trust, risk, and privacy concerns

Zarouali, B.; Strycharz, J.; Helberger, N.; de Vreese, C.

DOI
10.1080/0144929X.2021.2022208

Publication date
2022

Document Version
Final published version

Published in
Behaviour and Information Technology

License
CC BY-NC-ND

Citation for published version (APA):
Exploring people’s perceptions and support of data-driven technology in times of COVID-19: the role of trust, risk, and privacy concerns

Brahim Zarouali, Joanna Strycharz, Natali Helberger and Claes de Vreese

ABSTRACT
The COVID-19 pandemic has created one of the largest medical, financial, and social disruption in history. In the fight against this virus, many European governments have turned to collecting and using online data (for various technological applications) as a key strategic remedy. This study consists of data from a national representative survey in the Netherlands focusing on the extent to which data-driven technologies from the government can count on the support of the general public. By focusing on trust perceptions, risk beliefs and privacy concerns, we introduce a typology consisting of three subgroups: the sceptical, the carefree, and the neutral respondents. It was found that each of the three groups exhibit unique demographic characteristics. In addition, findings also revealed that these three identified groups have different support levels for specific digital solutions from the government. These findings contribute to an important and timely debate and entail relevant policy implications with regard to the democratic legitimisation of data-driven technologies in times of COVID-19.

1. Introduction

In the beginning of 2020, the world witnessed how the outbreak of a novel coronavirus (COVID-19) rapidly resulted in countries shutting down their borders and issuing very severe quarantines. In tackling the unprecedented challenges of this global pandemic, digital technology can play a pivotal role (UN 2020). Many European governments swiftly announced the possible collection and use of people’s personal data in the fight against the COVID-19 virus. Indeed, governments have started to collaborate with private companies, such as mobile network operators, location intelligence companies and AI technology companies to enrol data-driven tools in the fight against COVID-19 (Budd et al. 2020; UN 2020; Whitelaw et al. 2020). For a successful implementation of these digital tools, access to personal data (of citizens) is critical. However, much is yet to be explored about what people think of these digital solutions that exploit their data. As argued by the World Health Organisation, the effectiveness of such data-driven solutions largely depends on the levels of trust and support that a given population has in their own government to securely enrol this digital agenda (WHO 2020).

In addition, because the deployment of those technological measures typically comes along with what McDonald (2020) describes as ‘seizure of power’, and in so doing may affect individual rights (such as privacy and data protection), they require democratic legitimisation (Craglia et al. 2020). From the perspective of democratic legitimisation, it is important to investigate to what extent governmental data collection can count on the support of the people that will be affected by it (while focusing on whether certain parts of the population stand more behind this practice), as well as to explore to what extent they support specific technological measures as worthwhile solutions in the global pandemic.

In this study, a survey was distributed among a national representative sample of N = 907 respondents in the Netherlands. The aims were twofold. First, we wanted to explore people’s trust, risk and privacy concerns toward their data being collected by the government in the fight against COVID-19. The study draws on the tenets of social contract theory (Dunfee, Smith, and Ross 1999). Based on this framework, a first research question was formulated: which different groups of individuals can be distinguished based on trust perceptions, risk beliefs, and privacy concerns (in...
the context of collecting and using personal data in the fight against COVID-19? To answer this question, we used a cluster analysis (K-means) because it provides an effective and meaningful approach to classify the sampled individuals in our exploratory dataset into homogeneous groups and allows for a deeper understanding of these groups as well. In total, three groups were identified (i.e. the sceptical, the carefree and the neutral respondents), and their demographic characteristics and differences are discussed in more detail. The second research question of this paper is to further explore whether these subgroups have different levels of ‘support’ for specific digital measures in the combat against COVID-19. This resulted in a second research question: what are the differences between the subgroups in terms of support toward specific digital solutions of the government in the fight against COVID-19? We investigate people’s support toward a broad range of solutions including proximity tracing, hospital management, vaccine development, public communication, quarantine compliance, and disinformation detection (Budd et al. 2020; Gasser et al. 2020; UN 2020). We examine whether and to what extent the three subgroups which are identified (sceptics, carefree, and neutral) differ from each other when it comes to supporting these digital measures in the fight against COVID-19.

This paper makes multiple contributions. First, it examines how people differ in their perceptions toward the social contract in the context of governmental data collection. Measuring three constructs central to social contract in the context of collecting and using personal data in the fight against COVID-19. This resulted in a second research question: what are the differences between the subgroups in terms of support toward specific digital solutions of the government in the fight against COVID-19? We investigate people’s support toward a broad range of solutions including proximity tracing, hospital management, vaccine development, public communication, quarantine compliance, and disinformation detection (Budd et al. 2020; Gasser et al. 2020; UN 2020). We examine whether and to what extent the three subgroups which are identified (sceptics, carefree, and neutral) differ from each other when it comes to supporting these digital measures in the fight against COVID-19.

This paper makes multiple contributions. First, it examines how people differ in their perceptions toward the social contract in the context of governmental data collection. Measuring three constructs central to social contract in the context of collecting and using personal data in the fight against COVID-19. This resulted in a second research question: what are the differences between the subgroups in terms of support toward specific digital solutions of the government in the fight against COVID-19? We investigate people’s support toward a broad range of solutions including proximity tracing, hospital management, vaccine development, public communication, quarantine compliance, and disinformation detection (Budd et al. 2020; Gasser et al. 2020; UN 2020). We examine whether and to what extent the three subgroups which are identified (sceptics, carefree, and neutral) differ from each other when it comes to supporting these digital measures in the fight against COVID-19.

This paper makes multiple contributions. First, it examines how people differ in their perceptions toward the social contract in the context of governmental data collection. Measuring three constructs central to social contract in the context of collecting and using personal data in the fight against COVID-19. This resulted in a second research question: what are the differences between the subgroups in terms of support toward specific digital solutions of the government in the fight against COVID-19? We investigate people’s support toward a broad range of solutions including proximity tracing, hospital management, vaccine development, public communication, quarantine compliance, and disinformation detection (Budd et al. 2020; Gasser et al. 2020; UN 2020). We examine whether and to what extent the three subgroups which are identified (sceptics, carefree, and neutral) differ from each other when it comes to supporting these digital measures in the fight against COVID-19.

This paper makes multiple contributions. First, it examines how people differ in their perceptions toward the social contract in the context of governmental data collection. Measuring three constructs central to social contract in the context of collecting and using personal data in the fight against COVID-19. This resulted in a second research question: what are the differences between the subgroups in terms of support toward specific digital solutions of the government in the fight against COVID-19? We investigate people’s support toward a broad range of solutions including proximity tracing, hospital management, vaccine development, public communication, quarantine compliance, and disinformation detection (Budd et al. 2020; Gasser et al. 2020; UN 2020). We examine whether and to what extent the three subgroups which are identified (sceptics, carefree, and neutral) differ from each other when it comes to supporting these digital measures in the fight against COVID-19.

This paper makes multiple contributions. First, it examines how people differ in their perceptions toward the social contract in the context of governmental data collection. Measuring three constructs central to social contract in the context of collecting and using personal data in the fight against COVID-19. This resulted in a second research question: what are the differences between the subgroups in terms of support toward specific digital solutions of the government in the fight against COVID-19? We investigate people’s support toward a broad range of solutions including proximity tracing, hospital management, vaccine development, public communication, quarantine compliance, and disinformation detection (Budd et al. 2020; Gasser et al. 2020; UN 2020). We examine whether and to what extent the three subgroups which are identified (sceptics, carefree, and neutral) differ from each other when it comes to supporting these digital measures in the fight against COVID-19.

This paper makes multiple contributions. First, it examines how people differ in their perceptions toward the social contract in the context of governmental data collection. Measuring three constructs central to social contract in the context of collecting and using personal data in the fight against COVID-19. This resulted in a second research question: what are the differences between the subgroups in terms of support toward specific digital solutions of the government in the fight against COVID-19? We investigate people’s support toward a broad range of solutions including proximity tracing, hospital management, vaccine development, public communication, quarantine compliance, and disinformation detection (Budd et al. 2020; Gasser et al. 2020; UN 2020). We examine whether and to what extent the three subgroups which are identified (sceptics, carefree, and neutral) differ from each other when it comes to supporting these digital measures in the fight against COVID-19.

This paper makes multiple contributions. First, it examines how people differ in their perceptions toward the social contract in the context of governmental data collection. Measuring three constructs central to social contract in the context of collecting and using personal data in the fight against COVID-19. This resulted in a second research question: what are the differences between the subgroups in terms of support toward specific digital solutions of the government in the fight against COVID-19? We investigate people’s support toward a broad range of solutions including proximity tracing, hospital management, vaccine development, public communication, quarantine compliance, and disinformation detection (Budd et al. 2020; Gasser et al. 2020; UN 2020). We examine whether and to what extent the three subgroups which are identified (sceptics, carefree, and neutral) differ from each other when it comes to supporting these digital measures in the fight against COVID-19.

This paper makes multiple contributions. First, it examines how people differ in their perceptions toward the social contract in the context of governmental data collection. Measuring three constructs central to social contract in the context of collecting and using personal data in the fight against COVID-19. This resulted in a second research question: what are the differences between the subgroups in terms of support toward specific digital solutions of the government in the fight against COVID-19? We investigate people’s support toward a broad range of solutions including proximity tracing, hospital management, vaccine development, public communication, quarantine compliance, and disinformation detection (Budd et al. 2020; Gasser et al. 2020; UN 2020). We examine whether and to what extent the three subgroups which are identified (sceptics, carefree, and neutral) differ from each other when it comes to supporting these digital measures in the fight against COVID-19.

This paper makes multiple contributions. First, it examines how people differ in their perceptions toward the social contract in the context of governmental data collection. Measuring three constructs central to social contract in the context of collecting and using personal data in the fight against COVID-19. This resulted in a second research question: what are the differences between the subgroups in terms of support toward specific digital solutions of the government in the fight against COVID-19? We investigate people’s support toward a broad range of solutions including proximity tracing, hospital management, vaccine development, public communication, quarantine compliance, and disinformation detection (Budd et al. 2020; Gasser et al. 2020; UN 2020). We examine whether and to what extent the three subgroups which are identified (sceptics, carefree, and neutral) differ from each other when it comes to supporting these digital measures in the fight against COVID-19. From the perspective of democratic legitimacy, it is important to explore what extent governmental data collection (for the purpose of rolling out digital technologies) can count on the support or acceptance by the general population. To answer the question of public acceptance of digital solutions, prior research has often focused on three important factors: people’s trust beliefs, risk perceptions, and privacy concerns (Kalliopi 2016; Tang, Hu, and Smith 2008).

When it comes to trust, it is the result of a truster’s evaluation of how likely the trustee will behave according to the truster’s expectations (Baier 1986; Bauer 2014; Coleman 1990). In the context of this study, trust beliefs refer to degree to which people believe the government is trustworthy in protecting their personal data (Malhotra, Kim, and Agarwal 2004). It is important that citizens believe that their government is generally acting in their best interest and perceive their government to be capable of collecting and managing their personal data in a safe, secure and transparent way. Recent studies on technological solutions during COVID-19 (e.g. contact tracing apps) showed that trust in the government influences peoples’ response and perception toward that technological solution (Altmann et al. 2020; Moon 2020; Oldeweme et al. 2021).

Risk perception refers to people’s perceptions of the uncertainty and adverse consequences of handing over their data to the government (Dowling and Staelin 1994). Scholars have argued that perceived risk might be an important factor in determining citizens’ acceptance and adoption of online services from the government (Bélanger and Carter 2008; Warkentin et al. 2002), and that these perceptions can be an important barrier for accepting data collection by the government (Beldad, de Jong, and Steehouder 2011). In the context of Covid-19, it has been found that risk perceptions related to data security can influence the acceptance of digital solutions installed by the government, such as tracing apps (e.g. Albrecht et al. 2021).

Privacy concern refers to the degree to which a user is concerned about threats to their informational privacy online, usually in the context in which their personal data is being collected (Xu, Michael, and Chen 2013). Data privacy has always been a very important factor influencing citizens’ acceptance of governmental surveillance and monitoring (e.g. Nam 2019; Pavone and Esposti 2012; Thompson et al. 2008). In times of COVID-19, privacy concerns have been consistently found to be an important impediment in accepting and adopting technological solutions from the government (e.g. Abuhammad, Khabour, and Alzoubi 2020; Altmann et al. 2020; Chan and Saqib 2021; Kosterink et al. 2020). It is therefore one of the main
considerations to acknowledge in the technology-driven COVID-19 response (Redmiles 2021).

To connect these three concepts in a theoretical way, we will draw on the tenets of social contract theory in the next section.

2.2. Social contract theory

To explore people’s perceptions of the government collecting and monitoring personal data in the fight against the COVID-19 disease, we employ social contract theory (Dunfee, Smith, and Ross 1999). This theory has often been used as a meaningful conceptual framework to investigate people’s perceptions about data collection and informational privacy (e.g. Culnan and Bies 2003; Fogel and Nehmad 2009; Kruikemeier, Boerman, and Bol 2020; Malhotra, Kim, and Agarwal 2004; Martin 2016). As an important theory in political philosophy, it is relevant in explaining the legitimisation of governments to collect and process personal data of citizens (Economides 2018). Recent studies also highlighted the usefulness of a social contract perspective in times of COVID-19 (e.g. Ramelet 2020; Razavi et al. 1999; Voigt 2021), because the pandemic brings the implicit social contract between every citizen and their government to centre stage (Zivin and Sanders 2020). According to a social contract, all citizens are bound to do their share for the community by obeying the laws, rules and measures installed by the government during a pandemic; and in exchange, they benefit from the public goods provided by the government (e.g. safety, health, etc.), which are dependent upon citizen’s compliance with the rules and measures (Ramelet 2020).

In the context of the adoption of data-driven digital solutions, such a social contract could be understood as a ‘mutual agreement’ between individuals and the government about how their data is used and shared (Martin 2012, 2016) with the goal of managing the COVID-19 crisis. This means that individuals share their personal information with the government and trust that this institution handles their data safely because the government has the moral and legal obligation to protect it (i.e. the implied social contract) (Kruikemeier, Boerman, and Bol 2020; Okazaki, Li, and Hirose 2010). However, not every individual is willing to share personal data with the government as not everyone may perceive the social contract to be equally reliable and/or desirable (Martin 2016, 2012). Following this line of reasoning, an important question is: what are the ‘perceptions’ that people take into account in evaluating the implied social contract with the government? A recent study by Kruikemeier, Boerman, and Bol (2020) found that the three variables discussed in the previous section, i.e. trust beliefs, risk perceptions, and privacy concerns, play a pivotal role in a mutual social contract.

When it comes to trust and risk perceptions, it has been asserted that they provide important foundations for a social contract (Okazaki, Li, and Hirose 2010). A vast amount of research shows that these two variables, trust and risk, are most salient in social contract theory in the context of information sharing (e.g. Culnan and Bies 2003; Fogel and Nehmad 2009; Okazaki, Li, and Hirose 2010). More specifically, when individuals trust the government to handle their data safely, the individual perceives the social contract as more reliable; but when individuals see a high-level risk associated with the handling of their data by the government, the implied social contract becomes less reliable. In addition to these two variables, it was found that privacy concerns are inextricably linked to the implied social contract as well (Bansal, Zahedi, and Gefen 2010; Fogel and Nehmad 2009; Sheehan 2002). When people have the perception that their (informational) privacy is not respected and their data not kept safe, they can perceive social contract with the government as less reliable, while low privacy concerns are related to high reliability of social contract.

Taken together, Kruikemeier, Boerman, and Bol (2020) found that these three variables (i.e. privacy concerns, risk beliefs and trust perceptions) form the basis of a social contract; and that they can be used to cluster people in meaningful subgroups that are homogeneous with respect to these unobservable variables of interest. Following this evidence, we believe that these three variables (i.e. trust beliefs, risk perception and privacy concerns) would also be suitable to explore people’s perceptions with regards to the government’s intention to collect and use personal data to combat COVID-19. More precisely, we aim to investigate whether different subgroups of people can be identified based on these variables. In addition, we also aim to get a more fine-grained picture of these subgroups by exploring the similarities and differences between these clusters in terms of demographic composition. This allows to examine the prevalence and distribution of these subgroups among specific demographic cohorts. We propose the following exploratory research questions:

RQ1: (a) In the context of collecting and using personal data in the fight against COVID-19, which different groups of individuals can be distinguished based on trust perceptions, risk beliefs, and privacy concerns (b) what are the main demographic characteristics of the people in these subgroups?
2.3. Supporting data-driven technological measures

Having discussed people’s perceptions about the practice of governmental data collection, we will now zoom into how these data can be put to use to unlock the potential of technology in fighting the global pandemic (Whitelaw et al. 2020). That is, the reason why governments decide to collect and use data originating from a variety of digital sources is because it is a fundamental prerequisite for rolling out technological applications (Craglia et al. 2020). The use of such digital technologies has been presented as a key strategic remedy by governments in response to the COVID-19 pandemic. As argued by Gasser et al. (2020), technologies can be used to promptly alert and isolate exposed individuals (hereby preventing new infections), optimise public communication, increase the efficiency of healthcare, develop a vaccine, and improve quarantine measures.

This argument is entirely in line with recent reviews (e.g. Budd et al. 2020; Council of Europe 1990; Whitelaw et al. 2020), where some important areas were highlighted in which digital solutions can be harnessed: proximity tracing (i.e. using tracking technology to monitor individuals and inform them if they have been in contact with an infected patient), digital (public) communication (i.e. using digital communication platforms to quickly and efficiently inform the public), hospital management (using real-time data-driven methods to monitor the status of healthcare facilities, allocate health care resources, and increase hospital capacity), vaccine development (using recent advances in biotechnology and digital technologies to facilitate drug and vaccine development), and quarantine compliance (using digital technology to monitor whether people comply with quarantine and self-isolation measures) are important areas in which digital solutions can be harnessed.

In addition to these four digital solutions, the EU Science Hub (i.e. the European Commission’s science and knowledge service) also stressed the importance of an additional technological measure, i.e. using text mining algorithms and AI technology to automatically detect onlined disinformation about COVID-19 (European Commission 2020). The common denominator of these digital solutions is that they all use some kind of data about citizens (e.g. location data, patient health data, social media data, etc.). The effectiveness of these data-driven solutions is highly contingent on the support that citizens’ may vest in these solutions (WHO 2020). Therefore, this study aims to investigate the support among the population for the abovementioned six digital solutions.

As part of the implied social contract that we discussed earlier, several studies found that people’s privacy concerns, risk and trust perceptions (see RQ1) are foundational concepts that relate to how they respond to practices where their personal data is being collected online (Fogel and Nehmad 2009; Kruikemeier, Berman, and Bol 2020). In other words, these three concepts might be critical in determining people’s support level toward specific technological measures taken by their government. Following this line of reasoning, this study aims to explore to what extent the distinct subgroups from RQ1 (groups that are clustered based on their privacy concern, risk perceptions and trust beliefs) differ in terms of their support for the six digital solutions addressed above (i.e. digital communication, hospital management, vaccine development, quarantine, and disinformation) by the government in the combat against the coronavirus. Therefore, we propose this final research question:

**RQ2:** What are the differences between the subgroups (identified in RQ1) in terms of support toward specific digital solutions of the government in the fight against COVID-19?

### 3. Methodology

#### 3.1. Sample

The data consisted of a sample of \( N = 907 \) respondents in the Netherlands. This study was part of a larger survey focused on the impact of personalised communication and algorithms on society, which took on average 25 min to fill out. The response rate was 70%. The online survey ran from April 9 to April 20, 2020.
(11 days during the COVID-19 ‘lockdown’ period) and was distributed by IPSOS, a global market research firm. Stratification based on age, gender and level of education was applied to make the sample comparable to the general population. The participants had a mean age of 50.87 (SD = 15.90 years), and 46% of them were women. Around 26% had a low (no education or primary education), 50% moderate (secondary education), and 24% higher education level (bachelor, master or doctoral degree). A demographic overview of the sample is presented in Table 1.

Table 2. Constructs, items and their descriptive statistics.

<table>
<thead>
<tr>
<th>Constructs and items</th>
<th>M</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Trust perceptions</strong> (Cronbach’s alpha: 0.95)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>In the fight against COVID-19, the government has decided to collect and use online personal data from citizens. In this regard, to what extent do you agree with the following statements:</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1. The government is trustworthy in handling my personal data</td>
<td>4.19</td>
<td>1.57</td>
</tr>
<tr>
<td>2. The government tells the truth and fulfills its promises related to my personal data</td>
<td>4.15</td>
<td>1.53</td>
</tr>
<tr>
<td>3. I trust that the government keeps my best interest in mind when dealing with my personal data</td>
<td>4.49</td>
<td>1.64</td>
</tr>
<tr>
<td>4. The government is clear and consistent regarding the usage of my personal data</td>
<td>4.14</td>
<td>1.57</td>
</tr>
<tr>
<td>5. The government is honest with me when it comes to using my personal data</td>
<td>4.15</td>
<td>1.55</td>
</tr>
<tr>
<td><strong>Risk beliefs</strong> (Cronbach’s alpha: 0.92)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>In the fight against COVID-19, the government has decided to collect and use online personal data from citizens. In this regard, to what extent do you agree with the following statements:</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1. In general, it is risky to give my personal data to the government</td>
<td>4.48</td>
<td>1.59</td>
</tr>
<tr>
<td>2. There is a loss of privacy associated with giving my personal data to the government</td>
<td>4.73</td>
<td>1.59</td>
</tr>
<tr>
<td>3. There is too much uncertainty with giving my personal data to the government</td>
<td>4.72</td>
<td>1.53</td>
</tr>
<tr>
<td>4. Providing the government with my personal data can involve unexpected problems</td>
<td>4.55</td>
<td>1.58</td>
</tr>
<tr>
<td>5. I don’t feel safe giving my personal data to the government</td>
<td>4.21</td>
<td>1.68</td>
</tr>
<tr>
<td><strong>Privacy concern</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1. To what extent are you concerned about your online privacy when the government collects and uses your personal data in the fight against COVID-19?</td>
<td>4.24</td>
<td>1.69</td>
</tr>
<tr>
<td><strong>Support digital solutions</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>The government uses digital technologies in the fight against COVID-19. To what extent do you support the following digital solutions</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1. Using digital technology (e.g. apps) to track people who may be infected or infect others</td>
<td>4.11</td>
<td>1.77</td>
</tr>
<tr>
<td>2. The use of digital communication services such as WhatsApp by the government to quickly share information with the population</td>
<td>4.84</td>
<td>1.60</td>
</tr>
<tr>
<td>3. The use algorithms to determine which patients to prioritise in case of a shortage of healthcare capacity</td>
<td>3.40</td>
<td>1.79</td>
</tr>
<tr>
<td>4. Using Artificial Intelligence to find a drug/medicine against the coronavirus</td>
<td>5.15</td>
<td>1.49</td>
</tr>
<tr>
<td>5. Monitoring the use of mobile phones of citizens to check whether people are complying with the mandatory quarantine</td>
<td>3.82</td>
<td>1.83</td>
</tr>
<tr>
<td>6. Prevention of spreading disinformation about the coronavirus by means of automated analysis of social media posts</td>
<td>4.52</td>
<td>1.698</td>
</tr>
</tbody>
</table>

3.2. Measures

All items can be found in Table 2, along with their descriptive statistics. All instruments have been assessed on a 7-point scale from 1 (strongly disagree) to 7 (strongly agree). To measure trust perceptions, we used a validated instrument developed by Malhotra, Kim, and Agarwal (2004) consisting of five items. The items focus on the degree to which people believe that the government is trustworthy in handling and protecting their personal data. All items were averaged to form a single scale. Risk beliefs were assessed based on 7-point measure from the same authors (Malhotra, Kim, and Agarwal 2004). This scale consisted of five statements. The items focus on uncertainty and potential negative consequences of handing over personal data to the government. Scores on the five items were averaged to form a single scale for risk beliefs. Privacy concern was measured based on a single item, which has been adopted from prior studies (Chen and Chen 2015; Youn 2009). Respondents were asked to indicate the extent to which they were ‘concerned about their online privacy when the government collects and uses their personal data in the fight against Covid-19’. To measure respondents’ support for the digital solutions presented by the government, we asked respondents to indicate the extent to which they would support six different digital measures. These six solutions can be found in Table 2, and were discussed earlier in the theoretical framework.

3.3. Analytical strategy for clustering

To answer RQ1, a clustering method was used. Clustering is a commonly used exploratory data analysis technique to get a sense of the data structure. Although many clustering methods can be employed (e.g. hierarchical clustering, two-step clustering, etc.), we chose one of the most widely used cluster techniques: K-means clustering (Jain 2010). It refers to a robust and simple method to classify a sample of subjects based on a set of measured variables into a number of different groups or clusters so that similar subjects are placed in the same group (Cornish, 2007). The method starts with an initial guess for cluster centres, and each observation (i.e. respondent) is placed in the cluster to which it is closest. The cluster centres are then updated, and the entire process is repeated until the cluster centres no longer move (Charrad et al. 2014). A K-means cluster analysis is the recommended cluster solution when the dataset is moderate to large in size (e.g. large-scale surveys) (Brandtzæg, Heim, and Karahasanović 2011; Rogstad 2014).
In this analysis, $K$ denotes the number of clusters, which must be specified \textit{a priori} \cite{Jain2010,Pham2005,Ray1999}. To determine the number of clusters, there are many different methods \cite{Kodinariya2013}. In this study, we first started with two popular methods to visually inspect the clusters: the Silhouette method \cite{Rousseeuw1987} and the Gap Statistics method \cite{Tibshirani2001}. The R-package \textit{factoextra} \cite{Kassambara2016} was used for this, which provides data visualisations for these methods. After this visual inspection, we used the R-package \textit{NbClust} \cite{Charrad2014}, which computes 30 indices at once, in order to validate the optimal number of clusters. The results will be discussed in the next section.

### 4. Results

#### 4.1. Cluster analysis

To answer RQ1, we need to segment respondents into clusters (based on K-means clustering) of individuals with distinctly different perceptions based on trust, risk and privacy concerns (RQ1). The Silhouette method and the Gap Statistic method both suggest 3 clusters as the optimal number. To cross-check this, we inspected the 30 indices provided by the \textit{NbClust} R-package \cite{Charrad2014}. These indices revealed that 3 is the best number of clusters in our dataset (according to the ‘majority rule’). Therefore, we chose a three-cluster solution for our K-means cluster analysis. The identified clusters significantly differed in their trust perceptions, risk beliefs and privacy concerns (see Figure 1). Based on their scores on the three variables of interest, the three clusters are interpreted as (i) the sceptical respondents, (ii) the carefree respondents, and (iii) the neutral respondents. These groups will be discussed in the next paragraphs.

**4.1.1. Cluster 1: the sceptical respondents**

This first cluster accounts for 23% of the total sample ($N = 207$), which we named ‘sceptics’ because they stand out with their low level of perceived trust ($M = 2.22$), and considerably higher levels of risk beliefs ($M = 6.09$) and privacy concerns ($M = 6.01$). So out of the three clusters, this subgroup holds rather negative perceptions when it comes to the government collection their personal data in the fights against COVID-19. These respondents are characterised by a high level of trust perception ($M = 5.26$), and much lower levels risk beliefs ($M = 3.11$) and privacy concern ($M = 2.06$). The results of this subgroup tend to indicate rather positive perceptions toward tracking in COVID-19 times, and therefore, is somewhat opposite to the sceptic respondents.

**4.1.2. Cluster 2: the carefree respondents**

The second type of respondents that we identified are what we have called the carefree respondents, who account for around 29% of all respondents ($N = 267$).

**4.1.3. Cluster 3: the neutral respondents**

The final cluster represents the largest subgroup with around 48% of all respondents ($N = 433$). These respondents have ‘neutral scores’ for trust ($M = 4.55$), risk ($M = 4.67$) and privacy concerns ($M = 4.75$), and these scores are not very different from each other. So, these respondents have rather neutral perceptions toward the governmental data tracking practices, without holding strong positive or negative views.

#### 4.2. Cluster 3: the neutral respondents

The final cluster represents the largest subgroup with around 48% of all respondents ($N = 433$). These respondents have ‘neutral scores’ for trust ($M = 4.55$), risk ($M = 4.67$) and privacy concerns ($M = 4.75$), and these scores are not very different from each other. So, these respondents have rather neutral perceptions toward the governmental data tracking practices, without holding strong positive or negative views.

#### 4.3. Features of the respondents from the three clusters

Table 3 displays the distribution of several demographic variables in the three clusters. All estimates also contain a superscript that presents the results of the (two-sided) significance testing of the group differences. All tests were adjusted using the Bonferroni correction.

An important result that emerges from the data is the significantly higher proportion of respondents in the 18–24 age category among the sceptics, in combination with significantly lower proportion in the 55 + category. This seems to indicate a trend in which sceptical respondents are more often younger people. Gender does not seem to differ significantly among the three subgroups. As for education, an interesting pattern can be discerned from Table 3. On the one hand, respondents with neutral perceptions (indifferent) are significantly overrepresented in the ‘lower education’ category as compared to the representation of this education level in the two other clusters. On the other hand, the proportion of higher educated respondents is significantly higher among the sceptical user group as compared to the two other groups. So, this seems to suggest that higher educational degree is more likely to be associated with sceptical perception, whereas lower educational degree with indifferent perceptions.

#### 4.4. Support for technological measures of the government among the three clusters

The second research question (RQ2) was concerned with the relationship between the identified clusters and their support for digital solutions by the
government. Pairwise comparisons between the clusters with Bonferroni correction are presented in Table 3.

These results show that sceptical respondents tend to have little support for technological measures. In fact, the sceptic subgroup has significantly lower support for all technological measures compared to the other two groups, except for ‘using AI for vaccine development,’ where sceptics have a support level similar to the neutral group. A complete opposite pattern emerges from the data when it comes to the carefree respondents: they have significantly higher support for all digital solutions, except for one, i.e. using technology for monitoring quarantine compliance, where they have equal support levels as the neutral respondents. The neutral respondents can be positioned as somewhat in-between: they are more in support for digital solutions than the sceptical respondents, but less than the carefree people. So overall, these findings clearly indicate some meaningful differences between the three clusters, which will be further discussed in the general discussion.

### 5. General discussion

In this study, we first aimed to investigate how people evaluate the governmental practice of collecting data in the fight against the global pandemic. Drawing on the tenets of social contract theory (Dunfee, Smith, and Ross 1999), we investigated individuals’ trust, risk perceptions and privacy concerns, and by means of an exploratory cluster analysis of these factors introduced a typology consisting of three groups: the sceptical, the carefree, and the neutral respondents. It was found that each of the three groups exhibit unique characteristics. The sceptical respondents (23% of all respondents) are people with a high perceived risk and privacy concerns, in combination with a low degree of perceived trust when it comes to collecting their data. The carefree
respondents (29% of the respondents) have high levels of trust in governmental data collection, experience little risks, and have little privacy concerns. When we compare these two groups in terms of their composition, two interesting differences emerge. The group of sceptics has a significantly higher proportion of young people (18-34 years), as well as significantly higher educated people, whereas the carefree group consists of significantly more older people (55+years), as well as a significantly higher amount of lower educated people.

When it comes to the educational differences, the findings confirm prior literature that higher educated tend to have higher privacy concerns and lower levels of trust toward online data collection practices (Malhotra, Kim, and Agarwal 2004; Sheehan 2002). With respect to age group differences, our findings challenge the argument that younger adults are less sceptic toward online data collection, and generally less concerned about their privacy in comparison to older adults (since our dataset revealed the opposite) (e.g. Halperin and Dror 2016; Quan-Haase and Ho 2020; Van den Broeck, Poels, and Walrave 2015). A possible explanation for this could be ‘contextual integrity’ (Nissenbaum 2010), which states that perceptions of privacy violations might depend on the specific context or circumstances in which data flows take place. In the case of COVID-19, it has been shown that young people are less likely to develop severe symptoms as compared to elderly (Liu et al. 2020). Therefore, younger adults might perceive personal data collection as more drastic and privacy-infringing (because of the lower severity of the disease among these individuals) compared to older cohorts. In other words, the characteristics of this specific context could be the reason why different age categories make different risk-benefits assessments about the government collecting their data. However, this explanation did not directly flow from our data, so it should be a subject of future empirical testing that assesses the role of context and perceived severity in perceptions related to data collection.

In future scholarly endeavours, it is important to have a more detailed overview of the underlying reasons of why a significant part of the population (23%) is sceptical about governmental data collection, which was beyond the scope of the present study. These sceptic perceptions of data collection could be a great barrier to the widespread adoption of digital solutions during a pandemic. Understanding why people are sceptical about their governments collecting data for rolling out technological tools is crucial, because it can contribute to a broader uptake of these technological solutions (Zhang et al. 2020). When it comes to the neutral respondents, we found that a large group (48%) of people hold neutral levels of perceived trust, perceived trust, and privacy concerns with respect to governmental data collection. In terms of age distribution, this category is rather similar to the carefree group. However, when it comes to educational level, we see that this group of neutral respondents has significantly higher proportion of people with a lower education. Although it is important to explain this overrepresentation of lower educated people, it is difficult to do so because the ‘neutral’ group is rather ambiguous. It can be because they are ambivalent, undecided, unengaged, indifferent, or simply because they have not thought about it. From a scholarly perspective, it is important to obtain more fine-grained analyses to ‘unpack’ this ‘central tendency’ (i.e. neutral scores) to see whether there are additional variations in the responses of the public (e.g. ambivalence, see de Vries and Steenbergen 2013). From a societal perspective, with almost half of the population in this category (48% of the population), it places a great responsibility on governments to thoroughly understand these people when deploying new technologies and collecting data.

Personal data is at the heart of many digital solutions that European governments want to roll out in the fight against COVID-19 (Council of Europe 1990; Craglia et al. 2020). That is why this study also investigated to which extent the three identified groups support selected digital solutions presented by governments. On the one hand, we found that people with a sceptical mindset about the collection of their personal data by the government (i.e. sceptical subgroup) are more likely to have lower support for almost every technological solution on the digital agenda of the government. On the other, a carefree mindset (i.e. carefree subgroup) is associated with a higher level of support for the digital agenda of the government. So, people that evaluate governmental data collection in a neutral way when it comes to trust, risk and privacy, also tend to have moderate levels of support for specific digital solutions presented by that very same government. All in all, this suggests that the more certain people are in their social contract with the government (or lack of it), the more likely it is that this stance (positively or negatively) influences their support for the introduction of digital solutions by the government.

5.1. Theoretical and policy implications

From a theoretical perspective, this study makes an important contribution to the COVID-19 literature by showing that a social contract approach might help to understand and identify how people respond to collecting and using personal data in times of a pandemic.
Although other studies have already used this framework to explore people’s perceptions about data collection and informational privacy (e.g. Culnan and Bies 2003; Fogel and Nehmad 2009; Malhotra, Kim, and Agarwal 2004; Martin 2012), this study shows that a social contract perspective can also be interesting in studying people’s perceptions about how personal data is used and managed by the government with the goal of managing the COVID-19 crisis. In this respect, our results are also in line with the findings of Kruikemeier, Boerman, and Bol (2020), showing that trust beliefs, risk perceptions, and privacy concerns play a pivotal role in a mutual social contract. In addition, this study also contributes to the increasingly growing body of research on the technological responses to COVID-19 (e.g. Abuhammad, Khabour, and Alzoubi 2020; Albrecht et al. 2021; Lewandowsky et al. 2021; Oldeweme et al. 2021). Our findings confirm that perceptions about risk, privacy and trust are important considerations in accepting the digital solutions of a government during COVID-19; and in addition, can be used to create a typology of different subgroups within a population (i.e. the sceptical, the carefree, and the neutral individuals).

From a policy perspective, this study makes relevant contributions. Although data collection and usage for digital solutions are key in fighting the pandemic in an effective way, several concerns have recently been raised about the impact of these technologies on individual rights (e.g. privacy, data protection, personal autonomy) (Bengio et al. 2020). COVID-19 may have shifted the balance in data management and usage from individual rights towards the public good (Craglia et al. 2020). This has raised important questions about the democratic legitimacy of these data surveillance technologies (Craglia et al. 2020). As explained by Christensen and Lægreid (2020), the question of governance legitimacy is about citizens’ trust and support in their (democratic) government (in times of COVID-19), and includes concerns such as accountability and expectations. More precisely, it is about people trusting and supporting their government in deploying data-driven digital solutions. The contribution of this study is that it provides insights on the extent to which there is trust and support for governmental data collection and usage (i.e. the specific digital solutions), as well as information about which parts of the population that stand more behind these digital measures than others. Thanks to the cluster analysis, which provided an efficient segmentation approach that allowed to identify certain subgroups, policy makers can install targeted measures for these specific segments of the population.

In addition, one of the core debates in democratic legitimacy is the opinion of young people about their democratic government, with recent findings indicating a worrying increase in democratic dissatisfaction (see Foa et al. 2020). This might be a possible explanation why young people were found to be more sceptical and more concerned about their privacy with regards to governmental data collection (see ‘contextual integrity’ above). Therefore, it is important to keep track of and understand these dynamics in order to ensure youth legitimacy in times of a global pandemic. All in all, these findings could not only serve as a useful empirical basis for EU policy makers and regulatory bodies, but they also inform our understanding about the democratic legitimacy of these measures.

5.2. Limitations and future research
Finally, we also want to address some limitations that could inspire future research. First, this research is based on cross-sectional data. This means that it is not possible to articulate causal relationships. For instance, we found that trust, risk and privacy concerns are related to support levels of data-driven digital solutions, but we cannot claim that these perceptions directly enhance or undermine these digital solutions. Future research might therefore examine the causal structure of current findings. Second, data collection took place at one specific point-in-time, i.e. in the early days of the COVID-19 lockdown (April). It is not unlikely that perceptions in the context of data-driven technologies may change (in a positive or negative way) during the course of a crisis. Therefore, we also encourage scholars to explore whether current findings also hold at different points-in-time. Finally, as already mentioned earlier, we have a significant part of respondents that hold ‘neutral’ perceptions. Unfortunately, we do not have fine-grained data on why they scored neutral (it could be attributed to ambivalence, uncertainty, indifference, etc.; or, it can simply be a response bias whereby respondents avoid extreme answer categories). Future research might consider qualitative methods to obtain an in-depth understanding of people’s neutral stance with regards to technology in times of COVID-19.

Acknowledgements
This study was commissioned by the research initiative ‘Information, Communication & the Data Society’ (ICDS) from the University of Amsterdam.
Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

This work was funded by the Research Priority Area ‘Personalised Communication’ from the University of Amsterdam.

ORCID

Brahim Zarouali ORCID http://orcid.org/0000-0003-1042-4635
Claes de Vreese ORCID http://orcid.org/0000-0002-4962-1698

References


