How the level of detail of privacy regulation information regarding mobile applications affects trust and usage in times of a health crisis

Name author: Vera Veltman

Student ID number: 455451

University: Erasmus University Rotterdam

ERASMUS UNIVERSITY ROTTERDAM

Erasmus School of Economics

Bachelor Thesis Behavioural and Health Economics

How the level of detail of privacy regulation information regarding mobile applications affects trust and usage in times of a health crisis

Name student: Vera Veltman

Student ID number: 455451

Supervisor: Xiao Yu

Second assessor: Wen Qiang Toh

Date final version: 11th of July 2020

The views stated in this thesis are those of the author and not necessarily those of the supervisor, second assessor, Erasmus School of Economics or Erasmus University Rotterdam.

Abstract

This paper shows evidence from an experiment to answer the question of how the level of detail of information that is shared concerning privacy regulation affect trust regarding a COVID-19 tracker mobile application in times of the corona crisis. With the use of a survey, each respondent was assigned to one of the two treatment groups where the level of detail differed.

The results show there is no significant difference between the long or short versions of privacy disclosure in levels of trust and demand for usage of the mobile application. It is possible that due to the limitations of small sample size, the study is underpowered and thus the result is a false negative. Moreover, the coefficients of the long treatment between levels of trust and demand for usage are all negative. It may imply that trust and usage increases when given a less lengthy version of the same material. Based on these outcomes, recommendations for providers of a mobile application would be to keep provided information short when dealing with circumstances where a high level of trust and usage of mobile applications is needed. This can be done for example by using only bullet points when sharing information about the privacy regulation of those mobile applications.

Foreword

Dear reader,

This thesis research was written in a period in which there was a lot of uncertainty among the population. During the year 2020, the coronavirus has had a major impact on the daily lives of people around the world. Hopefully, this dissertation research has been able to contribute to future measures that are used with regard to technological options.

In addition, the situation of the coronavirus meant that people had to work at home in isolation which caused its own challenges. This also affected the way I interacted with my supervisor Xiao Yu. That is why I would like to thank her, besides all the help during the thesis period, for the flexible availability.

Kind regards, Vera Veltman

Table of contents

Abstract 2

Foreword 2

Introduction 4

Literature review 7

Information provision and trust 7

Level of trust in relation to privacy information in the usage of mobile/online applications 8

Methodology and Data 10

Experiment design 10

Data collection 12

Sample 13

Descriptive statistics 13

Analysis 15

Results 18

Conclusion and Discussion 25

Main conclusion 25

Limitations 26

Recommendations 27

Future research 27

Reference list 28

Appendix 31

Appendix A: Survey questions and information 31

Appendix B: Descriptive statistics of the control variables 35

Appendix C: Do-file STATA 37

Introduction

On January 11th, 2020 the first known death was reported of the coronavirus disease, or also called COVID-19 (Taylor, 2020a). The COVID-19 is an infectious disease caused by the coronavirus (World Health Organization: WHO, 2020). The virus mainly spreads from person to person by small droplets from the nose or mouth that are released when someone speaks, coughs, or sneezes. Symptoms differ from feverish to serious developments of difficulty in breathing and can even result in death. Mostly older people and those with underlying medical problems carry a higher risk to become seriously ill.

Because of the characteristics of COVID-19, it is very contagious amongst people. Its origin was thought to be in Wuhan, China around November 2019. Soon more cases were confirmed outside of China across the globe leading to a worldwide health crisis. By 30th of January 2020, the outbreak was already declared to be a Public Health Emergency of International concern ("Events as they happen", 2020). This led governments around the world to put in place all kinds of restrictions to prevent and reduce the spread of the virus. For example, Italy, which experienced a very steep growth in infected cases, undertook what is known as a full lock-down, preventing people from going outside when there is no necessity. This was done to 'flatten the curve', which refers to the expected outcome within countries without restrictions in place, which would cause many deaths and exceed the availability of hospitals and health centres (Buchholz, 2020).

However, these restrictive measures also have a downside. Not only does it cause disruption for people directly involved with the virus, but it also affects everyone else. "We are all affected by the coronavirus (COVID-19) pandemic, whether by the virus itself, government restrictions, or the economic uncertainty it is causing." (The Health Foundation, 2020, par.1). In order to relax these restrictions and diminish the negative effects it brings, while maintaining the much-needed advantages, other solutions came into play. As governments were being put under pressure (Matthers, 2020) multiple ideas were raised. On the 8th of April 2020, a report came out from the European Union which recommends the use of technology and data "in particular concerning mobile applications and the use of anonymised mobility data" (2020, p.7). Out of this recommendation and demand for a technological solution came a lot of initiatives for mobile applications (eHealth Network, 2020; Mingis, 2020). However, one of the requirements of the use of these mobile applications is a minimum level of users. For example, in The Netherlands a minimum of 60% of the population will have to make use of this application to make it beneficial (NOS, 2020). In the United Kingdom the minimum is even higher, with an aim of at least 80% of the population (Kelion, 2020). This is because, without this required amount, people would not get alarmed frequently enough after coming in contact with a positive tested person, such that the spread of the virus will still happen too rapidly.

In the process of getting to a level of sufficient mobile application users, the role of privacy regulation and information plays a big part (Gotink, 2020), especially as countries differ in how they share this information (Choudhury, 2020; Taylor, 2020b). Which brings the research question of this paper to the following:

How does the level of detail of information that is shared concerning privacy regulation affect trust with regard to a health beneficial app in times of a health crisis?

First of all, this research question raises the sub-question what the effect of the level of detail is on the level of trust. In addition, the importance of the number of mobile application users was previously mentioned. This also poses the sub-question of the relationship between the level of trust and the use of a mobile application.

Thus, in order to answer this research question and answer the sub-questions, several hypotheses are used. These are as follows:

Hypothesis 1: A more detailed explanation of privacy regulation information leads to a higher level of trust

Hypothesis 2: A more detailed explanation of privacy regulation information leads to a higher demand for usage of the mobile application

With the use of a survey two treatment groups were set up where both included a different amount of detailed explanation regarding the same privacy regulation information. The outcome after the analyses suggests that a more detailed version of the same set of information results in more distrust and lower demand for usage of the mobile application. The results of these statistical tests, however, showed a p-value much higher than the alpha of 0.05, indicating the outcomes are not significant.

The social and scientific relevance of this research is as follows. The social relevance of this paper is that its research is creating better insight into the effect of communication with regard to trust and usage of mobile applications. Currently, there are many concerns for potential users for some of the mobile applications which this research could be helpful for (Remeikis, 2020). Especially those with a public safety purpose like it is often used by governments. For example, in the situation of the COVID-19, it could even stop the coronavirus transmission (University of Oxford, 2020). In addition, the outcome of this research could provide mobile application providers with useful insights on how best to approach potential users to increase the number of users.

In terms of scientific relevance, multiple aspects within the research offer various new information and data as privacy policy is a relatively new development of the society and the research on this topic is relatively few, especially when it comes to privacy preferences on mobile applications (Miltgen & Smith, 2015). The existing studies have focused on either the preference of people by previous mobile application usage or trust in a technological form of information sharing. For example, Xu, Zhang, and Yan (2018) studied the previous use of mobile applications and the trust they have put in the agreement of the corresponding privacy regulations eventually linking this to recommendations in the appearance of future privacy regulations. While Feng et al. (2018) have performed studies on the trust people have with the information sharing technology called Mobile Crowd Sourcing (MCS). Moreover, when it comes to the behavioural aspects involved in these existing works, it focuses mainly on preference rather than the aspect of trust (Martin & Shilton, 2015).

The rest of the thesis is organized as follows, first, *Literature review* will provide in-depth analysis and highlight the information that is also used as a basis for the data retrieval in this paper. After this, data retrieval and statistical analysis are discussed in the M*ethodology*. The results obtained by these methods are then displayed and explained. Finally, the *Conclusion and Discussion* answer the posed research question and its hypotheses, followed by a discussion about the limitations across this research, recommendations that have come out of it, and future research proposals.

Literature review

This literature review will show an analysis of available material into the different parts that this research considers when answering the research question. The parts are based on the formulated research question and hypotheses.

Information provision and trust

In a previous study, evidence shows that people's intention to release personal information depends on trust, privacy issues of how the personal information is used, and information sensitivity (Bansal, Zahedi, & Gefen, 2010). It turned out that these are in themselves determined by personal characteristics. The study of Bansal et al. (2010) looked at providing personal health information related to online health services. Where the focus of the research is put on the trust people have regarding privacy issues. Linked to these concerns and trust are the personal characteristics which have an influence on this. Here they make a clear distinction between the information privacy, which is about having a sense of control over the information provided, and privacy issues, which is about the loss of privacy. The first, indicating active participation in the share of personal information, and the second indicating passive participation in the share of personal information. Bansal et al. (2010) establish personal characteristics as the basis for people's trust and concern. Where they make use of a person's choice preferences which depends on personal characteristics. Next, they conclude that when looking at a person's trust regarding the provision of personal health information, they should consider not only the circumstances but also the personal characteristics and experiences as the research shows that this impacts the perception of health information sensitivity in various ways. They state that “Emotional instability increases this sensitivity, whereas intellect marginally decreases, and agreeableness marginally increases, information sensitivity. Emotional instability heightens the fears of possible negative outcomes, which leads to increased sensitivity.” (Bansal et al., 2010, p.144).

Taddei & Contena (2013) studied the link between self-revealing behaviour and different types of trust. They measured trust in four aspects: towards the online social network providers, the legal framework, other users and perceived transparency about the information management. They found that when it comes to privacy concerns, this does not directly affect the self-revealing behaviour, where perceived control of information management and trust are essential factors that do affect the self-revealing behaviour. In this article, these essential factors are both recorded with the use of multiple measurement scales, namely, that of Perceived Control over Information (PCI) and General Trust (GT). Within this measurement of trust, the main distinction was made between the trust participants have in the provider and other users of the online social network. In another paper where the research of

Taddei and Contena (2013) relate to and is inspired by, again, a clear distinction is made in how it measures trust regarding other users of the software platform and the software platform provider (Krasnova, Spiekermann, Koroleva, & Hildebrand, 2010, p. 116).

Note that most research papers, like those listed above, and the sources they used, have adopted an approach that analyses a behavioural aspect of how this affects the provision of information. Or which also occurred, the other way around, where a behavioural aspect is analysed based on the type of information that is provided. All data on how the level of detail in information affects trust is scarce. Usually, this is because many of these research pieces follow the approach where software programs depend on the amount of information the user chooses to give. Especially when it comes to health-related information provision. However, if it is specifically a mobile application, the user must indicate whether to agree with the privacy regulation or not. The view on this order will then change and be discussed based on other articles in the following subchapter *Level of trust in relation to privacy information in the usage of mobile/online applications.*

Level of trust in relation to privacy information in the usage of mobile/online applications

Although mobile applications are part of an online network, it differs in many ways, both in function and in the device on which it is used. As a result, research documents looking at the broad focus of online activities related to the privacy information and trust would therefore not suffice with the specifications contained in this research.

Since mobile devices, let alone applications built for them, entered our lives not so long ago, the amount of research in this area is very limited. One of these few articles that have examined the combination of trust and privacy regulation information that is shown, has built a system based on the previous trust people have given away using their mobile device (Xu, Zhang, & Yan, 2018). This was then linked to a certain amount of private information appropriate to an application that then appears as a recommendation. While it does not explore what effect the privacy regulations have on trust, it does include the relationship of trust in the information privacy aspect of using mobile applications. This article also discusses and emphasizes that the privacy preservation issue is still open. And while the paper looks at how best to use the recommendation based on previous use, it did not consider people's awareness and how they experience the privacy information provided. Which still leaves the question open to how people respond to the extent of detail of privacy regulation information regarding trust.

The study by Feng et al. (2018) could shed more light on this part. In their paper, they discuss possible solutions for Mobile Crowd Sourcing (MCS) and examine the privacy, security and trust issues within this spectrum of technology. MCS is a form of crowdsourcing activities which are performed on smartphones or other mobile devices with the goal to either collect data actively or passively. When conducted passively this could lead to an intrusion of people's privacy which is an ongoing and active discussion when involved like the mobile application of the corona tracker. The research conducted by Feng et al. (2018) is done by analysing surveys that include these aspects. Among other things, they have found that privacy around identity is an issue due to the privacy measures that MCS can offer. How this information was presented to survey participants was not made clear. Though, the way in which trust was included was based mainly on the reliability, competence, availability and fairness of the MCS system. While trust itself was defined as "the confidence, belief, and expectation regarding the reliability, integrity, ability, and other characteristics of an entity" (Feng et al., 2018, p. 2975). This was then further subdivided into the different parts that MCS offers. It not only shows the approach to the analysis of the element of trust but also how trust is defined in the area of technological and online services. Since the situation, regarding the privacy in relation to a form of sourcing on mobile devices, is alike in this paper, these definitions and subdivisions of trust will also be applied in this study. Which means the subdivision of trust is split in the privacy measures of the application and the security of the provided information when using the application. This division aligns with the former focus of trust in the platform provider and is in addition to the previously mentioned Likert scale, which measures the level of this subdivision of trust.

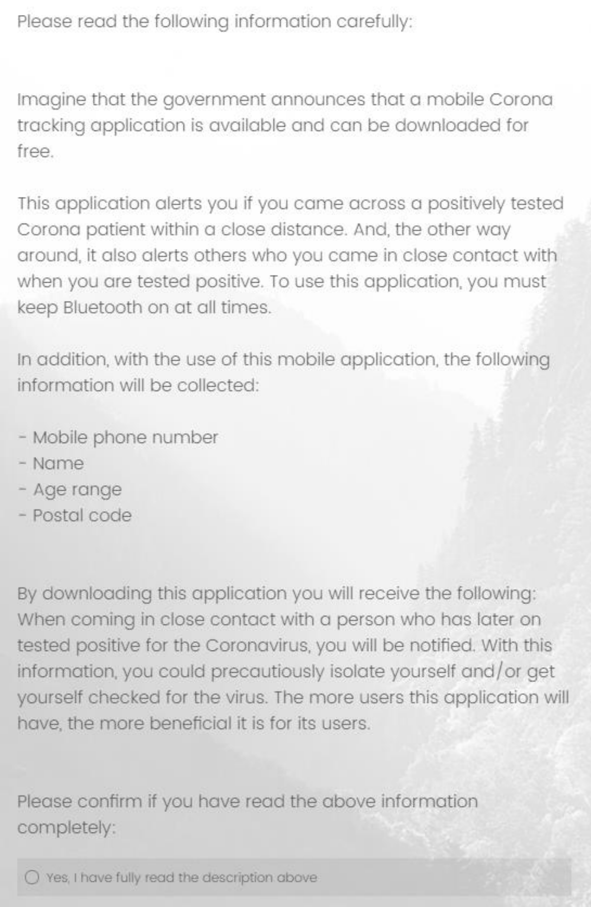
Methodology and Data

In this section the process to obtain the data and elements of the data will be discussed. First, the design of the experiment will be described and explained, then, the data collection, the sample of the dataset, the descriptive statistics, and lastly, the approach of the analysis in this research.

Experiment design

This research makes use of a survey to collect the required data. Due to restrictive measures that were in place during the period this research was conducted, no physical approaches were possible, only online options were available to gather data. Furthermore, with the questions in the online questionnaire both the essential data could be obtained, and a larger group could be reached compared to an individual approach.

For the survey, the software Qualtrics has been used. Within the design of the survey, the randomizer function was applied which made it possible to randomize the version a responder would receive.
As the respondent clicks on the survey link, a screen opens, starting with a short introduction, followed by either of the two treatments, the short or long description, of the mobile application is shown. Half of the subjects are randomly shown the long version and the rest are presented to the short version. Figure 1 and 2 show the screenshots of the long and short versions.



Please read the following information carefully:

Imagine that the government announces that a mobile Corona tracking application is available and can be downloaded for free.

This application alerts you if you came across a positively tested Corona patient within a close distance. And, the other way around, it also alerts others who you came in close contact with when you are tested positive. To use this application, you must keep Bluetooth on at all times.

In addition, with the use of this mobile application, the following information will be collected:

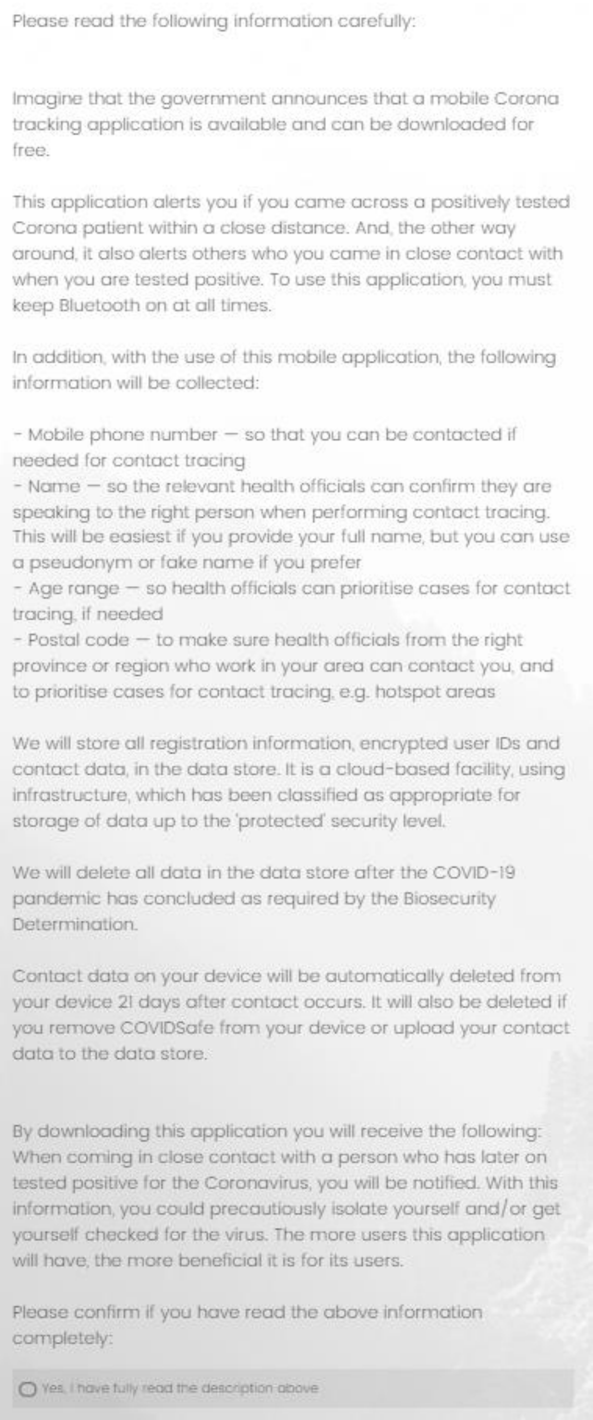
- Mobile phone number
- Name
- Age range
- Postal code

By downloading this application you will receive the following: When coming in close contact with a person who has later on tested positive for the Coronavirus, you will be notified. With this information, you could precautiously isolate yourself and/or get yourself checked for the virus. The more users this application will have, the more beneficial it is for its users.

Please confirm if you have read the above information completely:

○ Yes, I have fully read the description above

Figure 1. *Short description of the survey*



Please read the following information carefully:

Imagine that the government announces that a mobile Corona tracking application is available and can be downloaded for free.

This application alerts you if you came across a positively tested Corona patient within a close distance. And, the other way around, it also alerts others who you came in close contact with when you are tested positive. To use this application, you must keep Bluetooth on at all times.

In addition, with the use of this mobile application, the following information will be collected:

- Mobile phone number — so that you can be contacted if needed for contact tracing
- Name — so the relevant health officials can confirm they are speaking to the right person when performing contact tracing. This will be easiest if you provide your full name, but you can use a pseudonym or fake name if you prefer
- Age range — so health officials can prioritise cases for contact tracing, if needed
- Postal code — to make sure health officials from the right province or region who work in your area can contact you, and to prioritise cases for contact tracing, e.g. hotspot areas

We will store all registration information, encrypted user IDs and contact data, in the data store. It is a cloud-based facility, using infrastructure, which has been classified as appropriate for storage of data up to the 'protected' security level.

We will delete all data in the data store after the COVID-19 pandemic has concluded as required by the Biosecurity Determination.

Contact data on your device will be automatically deleted from your device 21 days after contact occurs. It will also be deleted if you remove COVIDSafe from your device or upload your contact data to the data store.

By downloading this application you will receive the following: When coming in close contact with a person who has later on tested positive for the Coronavirus, you will be notified. With this information, you could precautiously isolate yourself and/or get yourself checked for the virus. The more users this application will have, the more beneficial it is for its users.

Please confirm if you have read the above information completely:

○ Yes, I have fully read the description above

Figure 2. *Long description of the survey*

To make the description as close to reality as possible the content was based on an existing and live mobile application, where the concept of whether people choose to download and use the application is the same (Department of Health, Australian Government, 2020). As respondents will receive one of two versions of the survey, no control group is required in order to research differences between the two treatment groups. Respondents are then asked if they are willing to use the hypothetical mobile application or not with the question "Would you want to download and use this mobile application?"

with the possible answers of ‘Yes’ and ‘No’. They are also asked to answer a few questions in regard to the trust towards the mobile application with the statements “I ... that the security of the mobile application provider will prevent any leakages of my personal information.” and “I... the way the mobile application provider deals with the privacy of my personal information.” where the blank spots can be filled in by choosing from a range of 4 degrees of trust. Finally, the survey finishes off with some generic questions. The complete content of the survey and how this was shown to respondents can be found in *Appendix A*.

Data collection

The experiment was conducted from the 26th of May till the 31st of May 2020. Subjects were recruited primarily via the author's social network. In total 154 number of subjects have completed the survey, with 80 in the short version group and 74 in the long version group. However, due to a few reasons, a total number of 51 responses had to be taken out so that the data would be clean for analysis. One of the reasons why data got deleted was because respondents did not complete the survey by which the data could not be used. In total 35 got deleted because of this. Moreover, the preview version of the survey was filled in 7 times, which also needed to be taken out. Lastly, respondents who would spend less than 10 seconds on the page with the short description and less than 15 seconds on the long description also got deleted, as it is highly unlikely to have read the description within this amount of time and could lead to skipping of the treatment effect and could therefore lead to unreliable data. In total 9 responses were deleted because of this. Finally, a total of 103 responses were left with 53 in the short version group and 50 in the long version group. With the question of whether people would like to participate in this survey, no further (monetary) incentive was given other than the message of the help they would provide in contributing to the research.

Regarding the trust questions in the survey, as presented in *Appendix A*, the data could be collected to measure people' trust on both aspects of trust in privacy and security, based on the paper of Feng et al. (2018) which was discussed in the chapter of *Literature review*. These questions were shown immediately after the description of the privacy regulation measures of the corona tracker mobile application, together with the question of whether people would like to use this mobile application. Due to this setup, it makes it able to capture people’s first reaction after having obtained new information. Even though all the answers rely on self-reported data, which creates a possibility for inaccuracies within the dataset, it is the most feasible method to retrieve data.

Sample

When looking at the target group, the aim was to reach a group of respondents that varies across the different areas of the generic variables. However, the survey has been shared across online channels, where it reached a group of mainly friends and family. This limited the sought-after variety of the sample. In Table 1, *Appendix B*, the age of respondents has been grouped together to make a clearer overview of the whole dataset. In this table, it is shown that almost half of the respondents fall in the age group of between 18 and 25 years old. Furthermore, Table 4, *Appendix B*, shows that almost 80% has at least an educational level of that of a bachelor's degree or higher, which contributes to the non-variety in the dataset. In this same appendix, it can be seen in Table 3 that more than 90% is of Caucasian background. Looking at the gender spread across the respondents in Table 2, *Appendix B*, it appears that three quarters of the respondents are female. Altogether, this shows that variance across the different generic variables was unsuccessful.

Moreover, in order to base a conclusion on a large enough sample, sufficient data is required. The aim for the number of respondents was set on 30 people per treatment group, which would total in 60 respondents in total. Eventually, a total of 154 respondents participated in the survey of which 103 could be used after having this cleaned up. From this total amount, 53 are in the Treatment Short group and 50 in the Treatment Long group.

Descriptive statistics

The summary statistics of the cleaned dataset from this research are found in Table 5.

Table 5. *Summary statistics*

	Mean Treatment Short	SD Treatment Short	Mean Treatment Long	SD Treatment Long
Usage mobile application	0.510	0.505	0.444	0.503
Trust security	2.327	0.966	2.222	0.951
Trust privacy	2.388	0.975	2.289	0.895
N	53		50	

Note: SD = Standard Deviation. Usage mobile application is valued with 0 = Not use, 1 = Use, Trust security is valued with 1 = Do not trust at all, 2 = Distrust a little, 3 = Trust a little, 4 = Totally trust.

Table 5 shows that the mean trust and demand of usage are all higher in Treatment Short than in that of Treatment Long.

The independent variables obtained by the survey are all categorical, except for Age, which had an open response field where any number in years could be entered. Due to a very low number of answers for the category 'I would rather not say', namely 1, the categorical variable of Gender has been changed into a binary variable of Female, where 0 represents male and 'I would rather not say'. This binary change is also applied to the categorical variable of Ethnicity, where a very low number was from a Mixed or Asian ethnicity. This categoric variable is therefore changed into the binary variable of Caucasian. Table 6 shows the descriptive statistics of these independent variables and those of the time people have spent on the page with the description. Remarkable is that the average amount of time respondents spend on the short description page is longer than on that of the long description page.

Table 6. *Descriptive statistics: Baseline characteristics, Mann-Whitney U-test of independent variables between treatment groups, and time spend on both treatment group descriptions.*

	Mean Treatment Short	SD Treatment Short	Mean Treatment Long	SD Treatment Long	P-value
Age	33.184	13.838	34.244	15.718	0.646
Female	0.816	0.391	0.733	0.447	0.337
Caucasian	0.918	0.277	0.933	0.252	0.784
Education					
Highschool	0.082	0.277	0.044	0.208	0.464
College/MBO	0.163	0.373	0.133	0.344	0.686
Bachelor's degree or higher	0.755	0.434	0.822	0.387	0.430
Reading time	79.713	53.952	68.208	39.438	0.295

Note: SD = Standard Deviation. Female is valued with 0 = Male and unknown gender, 1 = Female, Caucasian is valued with 0 = Mixed or Asian, 1 = Caucasian, Reading time is in seconds.

The possible answers shown in the questions of gender and ethnicity are restricted as less as possible, such that everyone can find themselves in one of the options. For education this was done differently with purpose as the educational system can differ across the world and makes it harder to frame in strict categories. Thus, the main streams within education were displayed. Only the answer of college included a referral to studies in The Netherlands that overlaps the most with this variable, as the level of College might be otherwise incorrectly interpreted. Bachelor's degree and Highschool, on the other hand, are terms which overall are more frequently used internationally.

As seen in Table 6, the outcome of the Mann-Whitney U-test for all the independent variables suggests that the variables do not differ between the treatment groups.

The dependent variables are measured as follows: the respondents first get the description of the mobile application before receiving the question if they want to make use of this, followed by their level of trust in the security aspect of the application provider and that of the privacy aspect. The Likert scale was used to provide the answers for the respondents. This split in security and privacy, and the use of the Likert scale are based on the papers of Feng et al. (2018) and Krasnova et al. (2010), as is also discussed in the chapter *Literature review*.

Analysis

In order to answer the research question, and therefore first the hypotheses, multiple statistical analyses are required to answer these.

For the first hypothesis, which states that a more detailed explanation of privacy regulation information leads to a higher level of trust, the difference between the two treatment groups needs to be measured, regarding the trust in security and privacy. To get an image of the data distribution, first, a histogram is plotted for both the level of trust in security and privacy across both treatment groups including the kernel density line. This line approximates the density of the respondents from observations on the level of trust. This is also the function of the histogram itself, however, the line offers a clearer view whether this would be normal distributed data, which should be bell-shaped. As can be seen in Figure 3 and 4, all four graphs seem to have something like a peek in the middle and a descending slope towards the sides. However, these graphical analyses are not enough to assume whether the dataset is normally distributed. It does show a stronger skewness in Treatment Long towards distrust in security in Figure 3. Note that in both figures a higher outcome of trust has the meaning that there is more distrust.

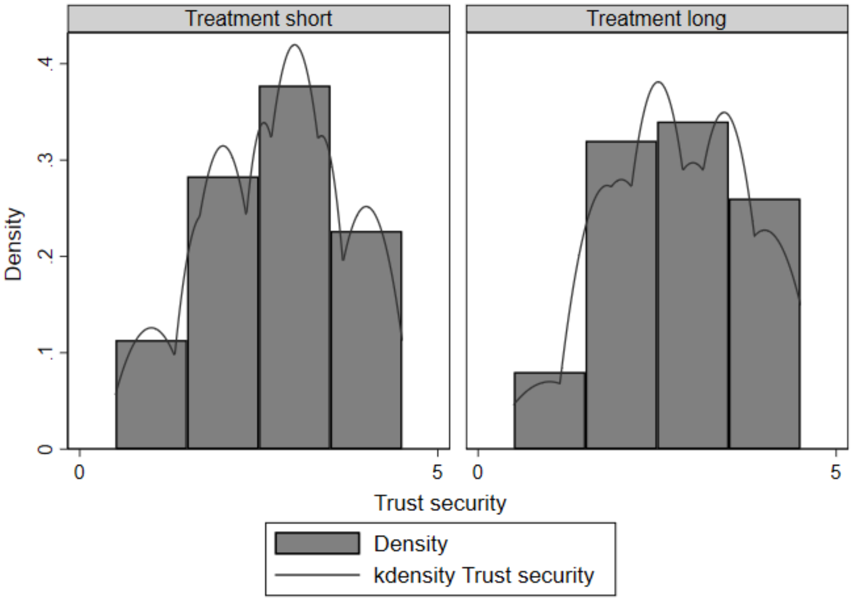


Figure 3. *Distribution level of trust regarding security over treatment groups*

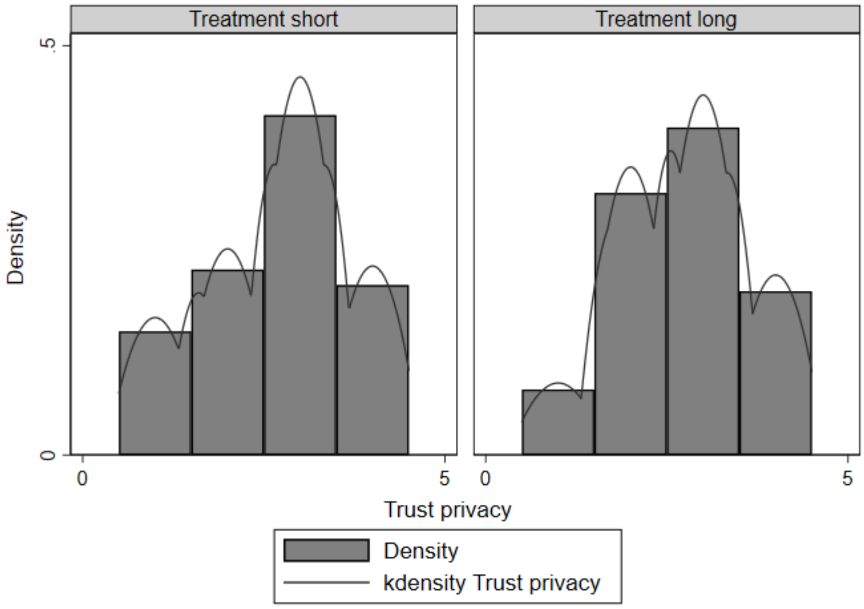


Figure 4. *Distribution level of trust regarding privacy over treatment groups*

Therefore, the analytical analysis which is used to check for this is a non-parametric method, namely the two-tailed Mann-Whitney U-test. Behind this test there is no assumption regarding the normality of the dataset. Moreover, the test is not sensitive for outliers which means no assumption regarding present outliers in the dataset must be fulfilled before using this test.

Additionally, a regression analysis is performed to see whether the variable of the treatment group affects the level of trust. Due to the categorical outcome and the order in this outcome of the level of trust, the use of an Ordered Logistic Regression will be made.

For the second hypothesis, which states that a more detailed explanation of privacy regulation information leads to a higher demand for usage of the mobile application, the difference between groups needs to be measured with regard to demand of usage where the Chi-squared test for independence will be used. This is also a non-parametric method which has no base assumption about the data distribution. Moreover, the test can be used for analysing nominal data, if the treatment groups are independent and equally exclusive. Regarding the regression analysis, the dependent variable in the regression of the second hypothesis is binary which changes the regression analysis performed, namely in this case the Probit Regression.

The specification of the regression analyses are as follows:

(1) Trust Security = b_0 + b_1 Treatment long + b_2 Female + b_3 Age + b_4 Caucasian + b_5 Education + α

(2) Trust Privacy = b_0 + b_1 Treatment long + b_2 Female + b_3 Age + b_4 Caucasian + b_5 Education + α

(3) Usage = b_0 + b_1 Treatment long + b_2 Female + b_3 Age + b_4 Caucasian + b_5 Education + α

Results

This chapter shows the multiple statistical analyses that have been performed in order to answer the hypotheses. It describes the results from these analyses and interpretation behind this.

For the statistical analyses used to answer both hypotheses, the software STATA has been used. As mentioned in *Methodology*, to test the first hypothesis a Mann-Whitney U-test is chosen to apply. For this hypothesis, this has been performed twice, first on the level of trust in security and second on the level of trust in privacy. The results overall show very high outcomes of the p-level across both levels of trust, as shown in Table 7. Therefore, the outcome of the Mann-Whitney U-test on trust does not give evidence stating both treatment groups would differ from each other in either that of trust in security as that of privacy.

Table 7*. Two-sample Mann-Whitney U-test and mean on the level of trust in security and privacy of the treatment groups*

	Mean Treatment Short	Mean Treatment Long	P-value
Trust security	2.327	2.222	0.618
Trust privacy	2.388	2.289	0.691
N	53	50	

Note: *Level of trust in security and privacy are valued with 1 = Do not trust at all, 2 = Distrust a little, 3 = Trust a little, 4 = Totally trust*

Table 7 also shows that in both levels of trust the group of Treatment Long have a lower mean outcome, stating a lower level of trust.

Graphically these outcomes are shown in Figure 5. These bar graphs show very little difference between both treatment groups for the mean trust in security and its standard deviation. Again, trust in privacy in Figure 6 shows very little difference between the means of the treatment groups. Note that a lower value is shown of the mean level of trust for Treatment Long in both Figure 5 and 6.

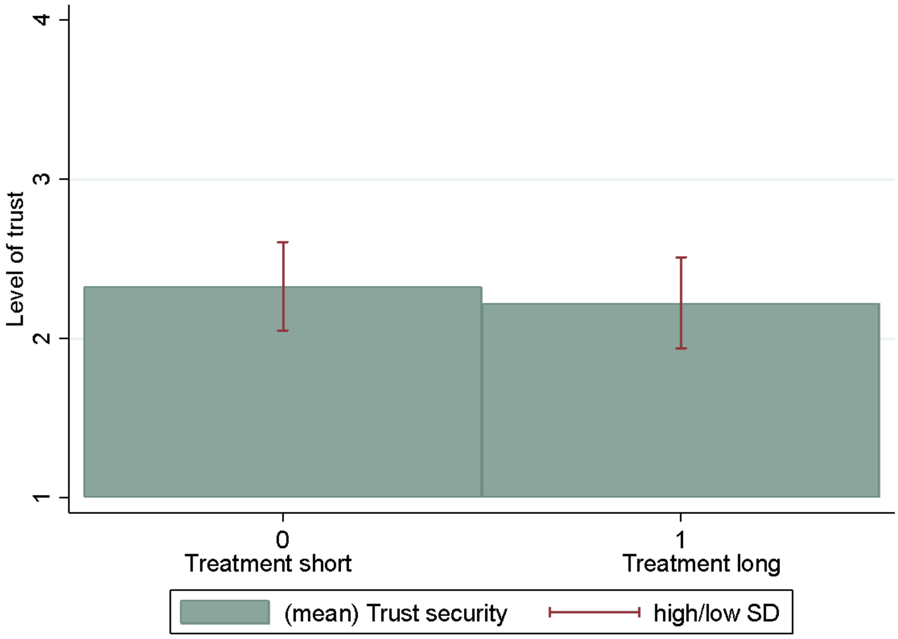


Figure 5. *Bar graphs of Level of trust in security over Treatment group with standard deviation (SD)*

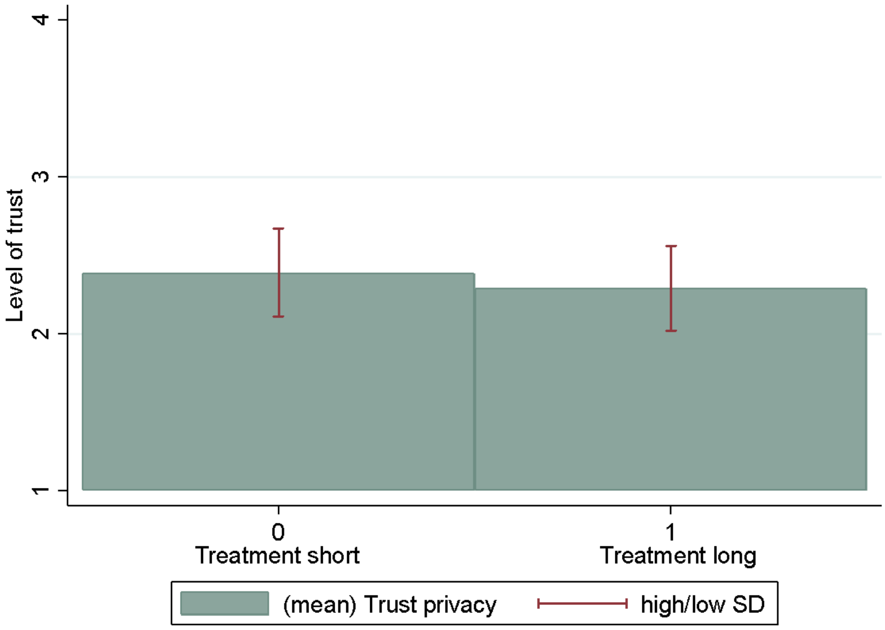


Figure 6. *Bar graphs of Level of trust in privacy over Treatment group with standard deviation (SD)*

In addition to the Mann-Whitney U-test an Ordered Logistic Regression has been performed. The variable of Treatment Long shows up insignificant, meaning there is not significant evidence suggesting the treatment group affects the outcome of the level of trust in security. However, its coefficient value shows more distrust in both security and privacy when provided with the long version of the regulation details, as is shown in both Table 8 and 9.

Table 8. *Regression Level of trust in security with control variables*

Variables	Level of trust in security
Treatment Long	-0.170
	(0.383)
Female	0.078
	(0.467)
Age	-0.046**
	(0.018)
Caucasian	-0.959
	(0.716)
Education	
College/MBO	1.680
	(1.273)
Bachelor's degree or higher	1.140
	(0.968)
Pseudo R^2	0.057

*Note: Standard errors are in parentheses. *p<0.10 **p<0.05 ***p<0.01. Level of trust in security is valued with 1 = Do not trust at all, 2 = Distrust a little, 3 = Trust a little, 4 = Totally trust.*

For the second regression stating the level of trust in privacy as the dependent variable, the main outcome seems to appear as in the previous regression shown in Table 8.

Table 9. *Regression Level of trust in privacy with control variables*

Variables	Level of trust in privacy
Treatment Long	-0.158
	(0.398)
Female	-0.102
	(0.501)
Age	-0.038**
	(0.018)
Caucasian	-0.584
	(0.496)
Education	
College/MBO	0.950
	(1.277)
Bachelor's degree or higher	0.486
	(0.909)
Pseudo R^2	0.033

*Note: Standard errors are in parentheses. *p<0.10 **p<0.05 ***p<0.01. Level of trust in privacy is valued with 1 = Do not trust at all, 2 = Distrust a little, 3 = Trust a little, 4 = Totally trust.*

Overall, the coefficient values of the treatment group suggest more distrust towards security and privacy when provided with the description of Treatment Long. Next to that, both regressions align as well in the coefficient directions regarding ethnicity. When someone is of another ethnicity than that of Caucasian, which was obtained in this dataset (see Table 4. *Appendix B*), the trust in both regressions increases, when all other variables stay constant. Lastly, for education level the trust increases strongest in both regressions when someone is of a College/MBO background.

For the second hypothesis the Chi-squared test is performed where the results are shown in Table 10.

Table 10. *Chi-squared test and mean on the demand of usage of the treatment groups*

	Mean Treatment short	Mean Treatment long	P-value
Usage	0.510	0.444	0.524
N	53	50	

*p-value<0.1, **p-value<0.05, ***p-value<0.01.

As shown in Table 10 the result of the p-value is higher than the significance level of 0.05, which gives evidence supporting both variable groups, Treatment groups and Usage, would be independent from each other.

Table 11. *Regression Usage with control variables*

Variables	Usage
Treatment group	-0.165
	(0.270)
Female	-0.069
	(0.340)
Age	-0.021*
	(0.011)
Caucasian	-0.513
	(0.509)
Education	
College/MBO	0.351
	(0.680)
Bachelor's degree or higher	0.369
	(0.583)
Constant	0.925
	(0.870)
Pseudo R^2	0.062

*Note: Standard errors are in parentheses. *p<0.10 **p<0.05 ***p<0.01. Usage is valued with 1 = Yes, 0 = No.*

The regression performed for the second hypothesis, namely the Probit regression in Table 11, shows an outcome alike of that of the first hypothesis. Note, however, that the coefficients cannot be interpreted as in linear regressions. Instead, a one-unit increase of the independent variable causes an increase in the size of the coefficient in the z-score of the dependent variable Usage. In this case, this would mean that an increase of age by one year would result in a decrease of the z-score of Usage by 0.021, keeping other variables in the regression constant. As well as the Treatment Long, which gives a decrease in z-value of 0.165, when other variables stay constant, which means it is less likely people provided with this version of Treatment are less likely to want to make use of the corona tracker mobile application.

In addition to the performed regressions, the interaction between variables has also been looked at. Due to the significant correlation of Age with the dependent variables of Trust security, Trust privacy, and Usage, the interaction between each of these variables have been tested. However, the results of these turned out all insignificantly, meaning the effect of one of the trust and usage variables does not depend on the Age variable and the other way around.

All the performed actions and tests in STATA stated in *Methodology* and *Results* are presented in the Do-file in *Appendix C*. How these results answer the previously posed hypotheses, and therefore the research question, are discussed in the chapter *Conclusion and Discussion*.

Conclusion and Discussion

In this section, the main conclusion will answer the research question and hypotheses. Next to that, the limitations are discussed that this research has come across. Thereafter, the recommendations based on the outcomes will be given and, finally, possible future research features will be discussed.

Main conclusion

Based on the retrieved results from this research, the research question '*How does the level of detail of information that is shared concerning privacy regulation affect trust with regard to a health beneficial app in times of a health crisis?*' can be answered with the help of the two phrased hypotheses.

First, hypothesis 1 states that a more detailed explanation of privacy regulation information leads to a higher level of trust. As was found in the results, the outcome of the Mann-Whitney U-test and regressions suggests otherwise. Even though the outcome suggests no difference between the two groups, the insignificant coefficient from both regressions suggest the opposite of the stated hypothesis, that a more detailed explanation of privacy regulation information leads to a lower level of trust. Therefore, there is not sufficient evidence to accept the first hypothesis.

Second, hypothesis 2 states a more detailed explanation of privacy regulation leads to a higher demand for usage of the mobile application. Results show, like for hypothesis 1, no outcome that suggests both treatment groups differ in usage. However, from the regression being performed, the coefficient suggests that when provided with a more detailed explanation of the same taken regulations, less demand for usage will occur. Thus, this leads to the outcome that there is no sufficient evidence to accept hypothesis 2.

The data retrieved, which shows that the mean reading time of Treatment Short is lower than that of Treatment Long, could also mean people spend less attention and time on long descriptions. This could perhaps also explain why people who are provided with more information have less trust and are less likely to use the mobile application. However, due to the unobserved environments the participants conducted the survey in, this cannot be said for certain.

Altogether, the research question can be answered that indeed the level of detail of information provided on privacy regulation negatively affects both trust and demand for usage.

Limitations

Although the randomization worked, the internal validity could be improved. Many variables have not been included in this study. Some have already been mentioned in the *Literature review*. Since not all variables are considered, this caused a biased result. Looking at the variables that are included, it only includes those internal variables that depend on the person, such as characteristic traits. In this research, this is the variable of trust and current demand for usage. However, there are not only many other internal factors, but those of the external factors are also important. As in this situation, they can affect the respondent's mood during the survey, for example, due to different weather conditions.

Also, the respondents read the description without any supervision or controlled environment. And even though a timer was put on the description page, it still does not offer assurance that the respondent actual read the text. Which would also affect the answers the respondent gives. Continuing the fact that respondents are not under supervision in any way, this could also mean that people fill in other generic aspects than would be the truth.

As is also shown in the descriptive statistics in the chapter *Methodology* the sample was not entirely spread over the different demographic variables, like ethnicity, gender, age, and education, decreasing its external validity. This causes that the results are not generalizable for the whole population but, in this research, mainly for the younger and higher educated females with a Caucasian ethnicity. One of the possible reasons for the distorted variety could be because of the limited channels that were used due to the corona restrictions, leaving only online options open to sharing the survey on.

To continue in the provided generic information, the question regarding education was offered with multiple answers. However, each of these answers could mean something else depending on where in the world you received the education.

Moreover, the survey is regarding a topic which is a lot in the news. Even though the randomization divides people who are very well informed and lesser informed on this topic across both treatment groups, this could still influence the levels of trust and usage itself or create a distortion in the data.

Finally, in this survey, only age was an open question whereas the rest was multiple-choice based on the Likert scale. This limits the respondent’s answer, where perhaps different more detailed results could be retrieved when applying a different form of answer possibilities.

Recommendations

With the outcome of the results, some recommendations can be made. First of all, for governments, as was pointed out in the *Introduction*, usage of the corona tracking applications is only beneficial with a certain number of users. As the results show that when more details are given, there will be more mistrust and a decrease in demand for use. It is therefore advisable, when providing detailed information about a mobile application, to keep this short. This can be done by, for example, showing the privacy regulation in a summarized way with the use of bullet points. It might also be interesting for those institutions willing to increase the number of mobile application users to look at the different aspects of information provision. However, more research is needed to specify what is furthermore required to achieve this higher rate of users.

Future research

There are multiple elements and options to work with for future research. First, the situation that is considered differs a lot compared to the regular situation people are in. Even though it has been said more epidemics will appear (Whiting, 2020), this will not be a constant situation throughout the years. The impact a crisis can have through cognitive heuristics (intellect), and emotional instability, makes for a possible different perception people have of information that is provided to them than they would have had outside of a crisis situation (Metzger and Flanagin, 2013; Dattilio, Freeman & Beck, 2007). It could be interesting to research the same elements but in a different situation, outside of a (health)crisis. And in line with this, more elements could be researched for besides the elements of privacy. Regulations contain a lot of information and it would be interesting to see what the response is when a non-privacy type of information is offered.

Also, next to using a bigger and more varied group than was used in this research, it would be interesting to add a third treatment group. The differences in the first and third treatment group could be made bigger. But with the addition of a middle group the nuance between the levels of detail of the information distribution would be possible to include in the research as well.

Finally, as is mentioned before, not all variables that could affect the level of trust and usage people want to make are included. Thus, more research in these missing variables is required. Possibilities for these variables are the different elements that the privacy regulations contain or the external elements which are present at the time of conducting the research. Also, the current usage of applications might show us how people treat the information about real-life usage against a self-reported usage demand.

Reference list

Bansal, G., Zahedi, F. M., & Gefen, D. (2010). The impact of personal dispositions on information sensitivity, privacy concern and trust in disclosing health information online. *Decision Support Systems*, *49*(2), 138–150. https://doi.org/10.1016/j.dss.2010.01.010

Buchholz, K. (2020, 21 March). *How fast is coronavirus spreading?* Retrieved on 30 April 2020, from https://www.weforum.org/agenda/2020/03/upward-trajectory-flattening-curve-how-countries-are-faring-coronavirus-covid-19/

Choudhury, S. R. (2020, 14 April). Apps collecting data to help stop the virus spread must limit sharing of information, cybersecurity expert says. Retrieved on 18 April 2020, from https://www.cnbc.com/2020/04/13/coronavirus-contact-tracing-technology-and-privacy-protection-for-apps.html

Dattilio, F. M., Freeman, A., & Beck, A. T. (2007). *Cognitive-Behavioral Strategies in Crisis Intervention, Third Edition* (3rd edition). New York, United States of America: Guilford Publications.

Department of Health, Australian Government. (2020, 27 May). *COVIDSafe app*. Retrieved on 14 May 2020, from https://www.health.gov.au/resources/apps-and-tools/covidsafe-app#about-the-app

eHealth Network. (2020). *Mobile applications to support contact tracing in the EU's fight against COVID-19* (1). Retrieved from https://ec.europa.eu/health/sites/health/files/ehealth/docs/covid-19_apps_en.pdf

European Union. (2020). *Official Journal of the European Union* (63). Retrieved from https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=OJ:L:2020:114:TOC

Events as they happen. (2020, 17 March). Retrieved on 30 April 2020, from https://www.who.int/emergencies/diseases/novel-coronavirus-2019/events-as-they-happen

Feng, W., Yan, Z., Zhang, H., Zeng, K., Xiao, Y., & Hou, Y. T. (2018). A Survey on Security, Privacy, and Trust in Mobile Crowdsourcing. *IEEE Internet of Things Journal*, *5*(4), 2971–2992. https://doi.org/10.1109/jiot.2017.2765699

Gotink, B. (2020, 18 April). *Nieuwe corona-app? Bouwers hameren op privacy, privacy en privacy*. Retrieved on 2 May 2020, from https://www.ad.nl/tech/nieuwe-corona-app-bouwers-hameren-op-privacy-privacy-en-privacy~a7ca4134/

Krasnova, H., Spiekermann, S., Koroleva, K., & Hildebrand, T. (2010). Online Social Networks: Why We Disclose. *Journal of Information Technology*, *25*(2), 109–125. https://doi.org/10.1057/jit.2010.6

Karreman, B., Werner, G., van der Molen, H., Osseweijer, E., Ackermann, M., Schmidt, H., & van der Wal, E. (2017). Academic Writing Skills for Economics and Business Administration. Amsterdam: Boom|Lemma

Kelion, L. (2020, 16 April). NHS coronavirus app to target 80% of smartphones. Retrieved on 18 April 2020, from https://www.bbc.com/news/technology-52294896

Martin, K., & Shilton, K. (2015). Why experience matters to privacy: How context-based experience moderates consumer privacy expectations for mobile applications. *Journal of the Association for Information Science and Technology*, *67*(8), 1871–1882. https://doi.org/10.1002/asi.23500

Matthers, D. (2020, 26 April). *Humanities experts are advising Germany on easing coronavirus restrictions*. Retrieved on 30 April 2020, from https://www.weforum.org/agenda/2020/04/german-humanities-scholars-enlisted-to-end-coronavirus-lockdown/

Metzger, M. J., & Flanagin, A. J. (2013). Credibility and trust of information in online environments: The use of cognitive heuristics. *Journal of Pragmatics*, *59*, 210–220. https://doi.org/10.1016/j.pragma.2013.07.012

Miltgen, C. L., & Smith, H. J. (2015). Exploring information privacy regulation, risks, trust, and behavior. Information & Management, 52(6), 741–759. https://doi.org/10.1016/j.im.2015.06.006

Mingis, K. (2020, 5 May). *Tech pitches in to fight COVID-19 pandemic*. Retrieved on 6 May 2020, van https://www.computerworld.com/article/3534478/tech-pitches-in-to-fight-covid-19-pandemic.html

NOS. (2020, 18 April). “Kans op succes corona-app klein”. Retrieved on 18 April 2020, from https://nos.nl/nieuwsuur/artikel/2330937-kans-op-succes-corona-app-klein.html

Remeikis, A. (2020, 27 April). *Covid safe: Australian government launches coronavirus tracing app amid lingering privacy concerns*. Retrieved on 2 May 2020, from https://www.theguardian.com/australia-news/2020/apr/26/australias-coronavirus-tracing-app-set-to-launch-today-despite-lingering-privacy-concerns

Taddei, S., & Contena, B. (2013). Privacy, trust and control: Which relationships with online self-disclosure? *Computers in Human Behavior*, *29*(3), 821–826. https://doi.org/10.1016/j.chb.2012.11.022

Taylor, D. B. (2020a, 28 April). *How the Coronavirus Pandemic Unfolded: a Timeline*. Retrieved on 30 April 2020, from https://www.nytimes.com/article/coronavirus-timeline.html

Taylor, J. (2020b, 21 April). Australia's coronavirus contact tracing app: what we know so far. Retrieved on 18 April 2020, from https://www.theguardian.com/world/2020/apr/17/australias-coronavirus-contact-tracing-app-what-we-know-so-far

The Health Foundation. (2020, 8 April). *What can we do to help those already facing disadvantage, in the COVID-19 outbreak?* Retrieved on 30 April 2020, from https://www.health.org.uk/newsletter-feature/what-can-we-do-to-help-those-already-facing-disadvantage-in-the-covid-19

University of Oxford. (2020, 16 April). *Digital contact tracing can slow or even stop coronavirus transmission and ease us out of lockdown | University of Oxford*. Retrieved on 2 May 2020, from http://www.ox.ac.uk/news/2020-04-16-digital-contact-tracing-can-slow-or-even-stop-coronavirus-transmission-and-ease-us

World Health Organization: WHO. (2020, 10 January). *Coronavirus*. Retrieved on 30 April 2020, from https://www.who.int/health-topics/coronavirus#tab=tab_1

Whiting, K. (2020, 4 March). *Coronavirus isn't an outlier, it's part of our interconnected viral age*. Retrieved on 27 May 2020, from https://www.weforum.org/agenda/2020/03/coronavirus-global-epidemics-health-pandemic-covid-19/

Xu, K., Zhang, W., & Yan, Z. (2018). A privacy-preserving mobile application recommender system based on trust evaluation. *Journal of Computational Science*, *26*, 87–107. https://doi.org/10.1016/j.jocs.2018.04.001

Appendix

Appendix A: Survey questions and information

Introduction – page 1

Welcome to this survey. My name is Vera and I am currently doing research for my bachelor thesis in economics at Erasmus University.
Your participation in this survey contributes to this and will **take about 3 minutes**. All your information and answers are **processed fully anonymously** and with care.

Your input is highly appreciated!

The following questions in this survey will be about your preference regarding the mobile corona tracking application.

Description, short version – page 2

Please read the following information carefully:
Imagine that the government announces that a mobile corona tracking application is available and can be downloaded for free.

This application alerts you if you came across a positively tested corona patient within a close distance. And, the other way around, it also alerts others who you came in close contact with when you are tested positive. To use this application, you must always keep Bluetooth on.

In addition, with the use of this mobile application, the following information will be collected:

- Mobile phone number
- Name
- Age range
- Postal code

By downloading this application, you will receive the following:

When coming in close contact with a person who has later on tested positive for the coronavirus, you will be notified. With this information, you could precautiously isolate yourself and/or get yourself checked for the virus. The more users this application will have, the more beneficial it is for its users.

Please confirm if you have read the above information completely:
[Yes, I have fully read the description above]

<u>Description, long version – page 2</u>
Please read the following information carefully:
Imagine that the government announces that a mobile corona tracking application is available and can be downloaded for free.

This application alerts you if you came across a positively tested corona patient within a close distance. And, the other way around, it also alerts others who you came in close contact with when you are tested positive. To use this application, you must always keep Bluetooth on.

In addition, with the use of this mobile application, the following information will be collected:

- Mobile phone number — so that you can be contacted if needed for contact tracing
- Name — so the relevant health officials can confirm they are speaking to the right person when performing contact tracing. This will be easiest if you provide your full name, but you can use a pseudonym or fake name if you prefer
- Age range — so health officials can prioritise cases for contact tracing, if needed
- Postal code — to make sure health officials from the right province or region who work in your area can contact you, and to prioritise cases for contact tracing, e.g. hotspot areas

We will store all registration information, encrypted user IDs and contact data, in the data store. It is a cloud-based facility, using infrastructure, which has been classified as appropriate for storage of data up to the ‘protected’ security level.

We will delete all data in the data store after the COVID-19 pandemic has concluded as required by the Biosecurity Determination.

Contact data on your device will be automatically deleted from your device 21 days after contact occurs. It will also be deleted if you remove COVIDSafe from your device or upload your contact data to the data store.

By downloading this application, you will receive the following:

When coming in close contact with a person who has later on tested positive for the coronavirus, you will be notified. With this information, you could precautiously isolate yourself and/or get yourself checked for the virus. The more users this application will have, the more beneficial it is for its users.

Please confirm if you have read the above information completely:
[Yes, I have fully read the description above]

<u>Usage and trust – page 3</u>
Imagine that the mobile application is now available for download for Android and IOS. Downloading is just one click away, for free, and barely takes any space on your storage.
Would you want to download and use this mobile application?
[Yes]
[No]

Taking in regard the information on the previous page about the mobile application, please fill in the blank spots in the following statements:

I ... that the security of the mobile application provider will prevent any leakages of my personal information.
[Totally trust]
[Trust a little]
[Distrust a little]
[Do not trust at all]

I... the way the mobile application provider deals with the privacy of my personal information.
[Totally trust]
[Trust a little]
[Distrust a little]
[Do not trust at all]

Generic questions – page 4

Finally, a few general questions.

What is your gender?

[Female]

[Male]

[Other]

[I would rather not say]

What is your current level of education?

[Less than highschool]

[Highschool]

[College/MBO]

[Bachelor's degree or higher]

What race or ethnicity do you relate to?

[Caucasian (European)]

[Latino/Hispanic]

[Middle Eastern]

[African]

[Caribbean]

[South Asian]

[East Asian]

[Mixed]

[Other, namely…]

What is your age (in years)?

[…]

Appendix B: Descriptive statistics of the control variables

Table 1. *Descriptive Statistics: Age per treatment group*

	Frequency Treatment Short	Percentage of Total	Frequency Treatment Long	Percentage of Total
Age <18	0	0.00%	1	0.97%
Age 18-25	27	26.21%	24	23.30%
Age 26-30	5	4.85%	4	3.88%
Age 31-40	10	9.71%	8	7.77%
Age 41-50	2	1.94%	2	1.94%
Age 51-65	8	7.77%	9	8.74%
Age >65	1	0.97%	2	1.94%
N	53		50	

Table 2. *Descriptive Statistics: Gender per treatment group*

	Frequency Treatment Short	Percentage of Total	Frequency Treatment Long	Percentage of Total
Female	42	40.78%	36	34.95%
Male	10	9.71%	14	13.59%
Rather not say	1	0.97%	0	0.00%
N	53		50	

Note: The categories with a frequency of zero are not counted in in this table. The category Other is therefore not included.

Table 3. *Descriptive Statistics: Ethnicity per treatment group*

	Frequency Treatment Short	Percentage of Total	Frequency Treatment Long	Percentage of Total
Caucasian	49	47.57%	47	45.63%
South Asian	0	0.00%	1	0.97%
East Asian	1	0.97%	1	0.97%
Mixed	3	2.91%	1	0.97%
N	53		50	

Note: The categories with a frequency of zero are not counted in in this table. The category Latino/Hispanic, Middle Eastern, African, Caribbean, and Other are therefore not included.

Table 4. *Descriptive statistics: Current level of education per treatment group*

	Frequency Treatment Short	Percentage of Total	Frequency Treatment Long	Percentage of Total
High School	4	3.88%	2	1.94%
College/MBO	8	7.77%	6	5.83%
Bachelor’s degree or higher	41	39.81%	42	40.78%
N	53		50	

Note: The categories with a frequency of zero are not counted in in this table. The category Below High School is therefore not included.

Appendix C: Do-file STATA

```
* Generate dummy variables
generate Female = 1
replace Female = 0 if Gender>1
generate Caucasian = 1
replace Caucasian = 0 if Ethnicity>1
gen Usage = 0
replace Usage = 1 if Use==1
generate Highschool = 1
replace Highschool = 0 if Education>2
generate College = 1
replace College = 0 if Education>3
replace College = 0 if Education<3
generate Bachelor = 1
replace Bachelor = 0 if Education<4
replace Bachelor = 0 if Education>4

* Descriptive statistics
tab Levelofdetail Age
tab Levelofdetail Gender
tab Levelofdetail Education
tab Levelofdetail Ethnicity

summ Age if Levelofdetail==0
summ Age if Levelofdetail==1
ranksum Age, by (Levelofdetail)

summ Female if Levelofdetail==0
summ Female if Levelofdetail==1
ranksum Female, by (Levelofdetail)

summ Caucasian if Levelofdetail==0
summ Caucasian if Levelofdetail==1
ranksum Caucasian, by (Levelofdetail)
```

```
summ Highschool if Levelofdetail==0
summ Highschool if Levelofdetail==1
ranksum Highschool, by (Levelofdetail)

summ College if Levelofdetail==0
summ College if Levelofdetail==1
ranksum College, by (Levelofdetail)

summ Bachelor if Levelofdetail==0
summ Bachelor if Levelofdetail==1
ranksum Bachelor, by (Levelofdetail)

summ Usage if Levelofdetail==0
summ Usage if Levelofdetail==1
summ Trustsecurity if Levelofdetail==0
summ Trustsecurity if Levelofdetail==1
summ Trustprivacy if Levelofdetail==0
summ Trustprivacy if Levelofdetail==1
summ PageSubmit if Levelofdetail==0
summ PageSubmit if Levelofdetail==1
ranksum PageSubmit, by (Levelofdetail)

* Mann Whitney U-test for hypothesis 1
ranksum Trustsecurity, by (Levelofdetail)
ranksum Trustprivacy, by (Levelofdetail)

* Regression for hypothesis 1
ologit Trustsecurity Levelofdetail Female i.Education Caucasian Age, robust
ologit Trustprivacy Levelofdetail Female i.Education Caucasian Age, robust

* Chi-square test for hypothesis 2
tab Usage Levelofdetail, chi2

* Regression for hypothesis 2
probit Usage Levelofdetail Female i.Education Caucasian Age, robust
```

```
* Bardiagram (mean) with Standard Deviation
collapse (mean) meantrusts= Trustsecurity (sd) sdtrusts=Trustsecurity (count) n=Trustsecurity ,
by(Levelofdetail)
generate hitrusts = meantrusts + invttail(n-1,0.025)*(sdtrusts / sqrt(n))
generate lowtrusts = meantrusts - invttail(n-1,0.025)*(sdtrusts / sqrt(n))
twoway (bar meantrusts Levelofdetail) (rcap hitrusts lowtrusts Levelofdetail)

collapse (mean) meantrustp= Trustprivacy (sd) sdtrustp=Trustprivacy (count) n=Trustprivacy ,
by(Levelofdetail)
generate hitrustp = meantrustp + invttail(n-1,0.025)*(sdtrustp / sqrt(n))
generate lowtrustp = meantrustp - invttail(n-1,0.025)*(sdtrustp / sqrt(n))
twoway (bar meantrustp Levelofdetail) (rcap hitrustp lowtrustp Levelofdetail)

* Regression level of trust on Usage with all variables as control variables
regress Usage Trustsecurity Trustprivacy Age Female i.Education Caucasian Levelofdetail, robust

* Interaction check
ologit Trustsecurity Levelofdetail##c.Age Female i.Education Caucasian, robust
ologit Trustprivacy Levelofdetail##c.Age Female i.Education Caucasian, robust
ologit Usage Levelofdetail##c.Age Female i.Education Caucasian, robust
```