Trusting Technology -
An Exploratory Study on the Influence of Trust on Memory

by

Maria Christine Genée
S1897942

Bachelor Thesis
Submitted to the Department of Human Factors and Engineering (HFE)
Faculty of Behavioural, Management and Social Sciences (BMS)
At the University of Twente
In Partial Fulfilment of the Requirements
For the Degree of Bachelor of Science in Psychology

1st Supervisor: Dr. S. Borsci
2nd Supervisor: Dr. R.H.J. Van Der Lubbe

July 2020
Enschede, The Netherlands

Abstract

This exploratory study investigates differences in people's detection ability of untrustworthy and trustworthy devices as well as other stimuli and if memory ability and predisposition to trust technology influences this. The data of 83 participants was collected, who viewed pictures of faces, scenes and devices taken from standardised databases and filled out the McKnight Trust in Technology scale. Using signal detection theory, the ability of participants to discriminate among target stimuli and new stimuli was calculated. The resulting discriminability indexes (d') of all stimuli and type of stimuli were used to establish detector groups, which categorised participants as either better detectors of trustworthy or untrustworthy stimuli. Results showed that the ability to recognise stimuli was influenced by memory ability, though no substantial difference in people's ability to recognise either untrustworthy and trustworthy stimuli was found. Moreover, no influence of the level of trust in technology in the detection groups was found. However, a tendency seems to be evident between the general level of trust in technology and the ability to better recognise untrustworthy devices.

Introduction

The cornerstone of the stability of society, in the long run, is trust (Alesina & La Ferrara, 2002; Hosmer, 1995; Baba, 1999). It is trust that makes interactions between individuals or groups effortless and more effective, for this to happen, mutual trust from both parties is necessary (Knack & Keefer, 1997). With the exponential growth in technology and the dawn of artificial intelligence in recent years, research studies on trust have taken on entirely new proportions.

From a psychological point of view, the term trust is often defined as the willingness of an individual to depend on another on account of the characteristics of this other party (Rousseau, Sitkin, Burt, & Camerer, 1998). It highlights the fact that people have expectations regarding performance and actions of others. Research in this field has also focussed on investigating trust's 'darker side', namely distrust. The term has been defined as, first and foremost, the absence of trust and secondly, as having the expectation of our safety being endangered by any sort of actor (Kramer, 1999). Further research has defined it as an actor's inclination to impute dangerous intentions (Lewicki, McAllister, & Bies, 1998). Moreover, previous works suggest that distrust is formed via experiences or interactions with others (Frederiksen, 2011).

All forms of trust play a vital role when investigating the adoption of technologies, yet literature on the relationship between untrust and technology is scarce. In this context, untrust lies between trust and distrust. It is defined as a measure of how small the amount of faith the truster places in the trustee. To clarify, untrust may be seen as a form of positive trust yet lacks just enough to not cooperate (Marsh & Dibben, 2005). The present study focuses on trust prior to the use of technology. As trust and distrust are formed through experience, this study focusses on untrust as a concept. Untrust is not based on experience and is thus more suitable and measurable in regard to the intention to adopt technology.

However, the mere existence of trust in human-technology interaction has been called into question. Some have contended that trust cannot occur between a human and a device. Friedman, Peter, Khan, and Howe (2000) hold to their claim that people only trust other people and not technology. Luhmann (1979) states that since an emotional bond cannot be established between a human and technology, trust in this interaction relies solely on a "presentational base".

In contrast, many researchers claim the opposite: studies have confirmed that trust between humans and technology is essential and determines how people choose to interact with various technologies. A study on the use of automated technology by Muir and Muray (1996) found that trust influences the user to adopt automated technology. A lack of this can damage the intention to use these systems (McKnight, Choudhury, & Kacmar, 2002). The Technology Acceptance Model (TAM) (Davis, 1989) describes how users accept technology, breaking the intention to use down to two components - perceived ease of use and perceived usefulness. Another factor that has been researched regards system quality, more precisely its visual appeal and navigational structure. These two components of system quality predicted the level of trust users place in technology (Vance, Elie-Dit-Cosaque, & Straub, 2008). Research has also shown that trust toward technology is also influenced by the amount of human features displayed by that system (Lankton, McKnight, & Tripp, 2015). Using studies on trust, trust toward systems (TTS) was conceptualised and is said to be shaped throughout the interaction with a given system. TTS exists even before the interaction with a system based on expectations (Borsci, Buckle, Walne, & Salanitri, 2018) - an essential component in the present study.

Another key concept in the present exploratory study of (un)trust prior to the use of technology is the predictive cheater mechanism module (Verplaetse, Vanneste, & Braeckman, 2007). It supports people in detecting how cooperative another party will be. Results showed that the participants could successfully distinguish faces of non-co-operators from co-operators when viewing pictures. Moreover, studies have confirmed the existence of this adaptive skill of remembering breaking of social norms and deceivers better (Cosmides, 1989; Comides & Tooby, 1992; Chiappe & Brown, 2004; Mealy, Daood, & Krage, 1996; Oda, 1997). However, these studies do not investigate what factors influence this cheater detection mechanism. The present study attempts to build upon these previous studies on cheating detection to explore if people can discriminate untrustworthy stimuli better compared to trustworthy stimuli.

In addition to the adaptive ability to recognise cheaters, memory and the type of stimulus seem to influence the recognition of (un)trustworthiness. Research has shown that humans have a general tendency to remember potential threats better than situations that are apparently harmless (McBride, Thomas, & Zimmerman, 2013; Hou, & Liu, 2019). The way in which people perceive their surroundings is based on both the incoming sensory input at that moment and their existing knowledge. This allows them to form expectations and, in turn, influences the overall recognition of the stimulus (Bar, 2007). These evolutionary mechanisms are not only evident in the perception of situations in general, but also when looking at faces (Kroneisen, 2018; Wendt, Weymar, Junge, Hamm & Lishke, 2019; Rule, Slepian, & Ambady, 2012; Mattarozzi, Todorov, & Codispoti, 2015). A study utilising event-related potentials (ERPs) on a memory task, found that after one week, participants recalled untrustworthy faces easier than trustworthy ones. People tend to have an advantage in long-term memory that allows them to recognise untrustworthy faces better (Weymar, Ventura-Bort, Wendt, & Lischke, 2019). Moreover, perception, learning and memory processes are affected by perceived untrustworthiness (FeldmanHall et al., 2018; Wendt et al., 2019). Encoding and retrieval of a stimulus are also affected by the perceived trustworthiness of a face (Mattarozzi et al., 2015). However, how technologies affect our trust judgement and level of general trust has yet to be investigated. The present study investigates whether these survival mechanisms are evident when viewing technologies before interacting with them and therefore incorporates all three stimulus types (faces, scenes and devices) to dissect if interactions occur between these stimuli and the general level of trust. Given the research mentioned above, the present study investigates whether untrust would be better recognised and remembered.

In the present study, participants need to make a decision if the given stimulus has been shown before or not. Signal detection theory (SDT) presumes participants performance is never perfect and illustrates their discriminatory ability in the context of uncertainty (Macmillan & Creelman, 2004). Thusly, SDT is applied here to measure the participant’s performance in recognition of the stimuli presented. Hence, it obtains a participants sensitivity in the discriminatory ability of old and new stimuli in the present study. From these discriminability indexes (d’), detector groups were created to categorise better untrustworthy detectors and better trustworthy detectors, on the basis of how many correctly identified old stimuli were recognised and whether these were trustworthy or not.

Furthermore, many findings of offline trust studies seem applicable to the online world (McKnight, 2005; Sasse, 2005). Due to these similarities, some measurements tend to humanise technology by using some of the same constructs as in trust measurements of human-human interaction, for example, benevolence, competence, ability and integrity (Vance et al., 2008; Wang & Benbasat, 2005). However, many technologies lack certain human-like functions, such as being able to make ethical decisions (Lankton et al., 2015). Therefore, in order to be able to explain the level of trust that exists in human-technology interaction, constructs like helpfulness, functionality, and reliability of a given technology are used (McKnight, Carter, Thatcher, & Clay, 2011). The present study builds on studies of offline trust and explores if the general level of trust in technology could explain a person’s enhanced untrustworthy detection ability.

To summarise, the main aims of the study are, first, to examine whether stimuli, that has been categorised as either trustworthy or untrustworthy in previous studies, are recognised better by taking memory ability and type of stimuli into account. Tested is how well participants can recognise old stimuli amongst new. Secondly, the present study investigates if this detection ability is also linked to the person's predisposition of trust in technology.

Methods

Design

A correlational design was employed. In doing so, the influence of memory ability measured during the pre-trial could be compared to the participant's scores of recognition of old and new stimuli. Additionally, the level of trust in technology was compared to these recognition scores as well.

Participants

The study consisted of 83 participants, which were recruited via convenience sampling. Participation in the study was voluntary. The requirements for partition was sufficient proficiency in the English language. People with mental or neurological deficits (e.g., inability to recognise faces, objects), have a photographic memory, and/or are under the age of 18 were excluded from participation. Participation requirements and exclusion criteria were asked in the consent form. The mean age within the study was 28.35 (SD=11.931, min= 19, max= 73). In addition to this, 42 females and 41 males took part. The most frequent completed level of education was a Bachelor's degree.

Materials

Using Qualtrics, the informed consent form for participation was created. Additionally, a questionnaire was designed using this software. The questionnaire collected demographic data of the participants (see Appendix A). Furthermore, the seven-item McKnight questionnaire for Trust in Technology (McKnight et al., 2011) was incorporated on a five-point Likert scale - from Faith in General Technology four items were used, and three items were taken from the Trusting Stance – General Technology (McKnight et al. 2002) (see Appendix B).

The experiment was designed using PsychoPy3. It consisted of 80 images which were derived from two standardised datasets. Images were taken from the Chicago Face Database from which 20 faces, who had been identified as defectors (rated less trustworthy) and 20 faces identified as cooperators (rated more trustworthy) (Ma, Correll, & Wittenbrink, 2015). Furthermore, 40 scenes had been selected from the socio-moral database SMID, where images had been identified as fair and moral (20 images) or unfair and immoral (20 images) (Crone, Bode, Murawski, & Laham, 2018). Lastly, 40 images of devices were discerned, which, based on consumer reviews and reliability of the product, were grouped together based on their trustworthiness (20 trustworthy, 20 untrustworthy). Devices categorised as trustworthy were on the market from 2017 onward and had no reported risks in usage or issues. Selected untrustworthy devices were reported as problematic and potentially endangered people's lives (CPSC.gov., n.d.) (see Appendix C).

Procedure

Each participant agreed online to the consent form before beginning the experiment. The experiment consisted of two stages – the pre-trial and the trial. In the first phase (pre-trial), participants viewed 20 images of flags. Every participant was instructed to attentively look at the stimuli presented. Each image was shown on the screen for three seconds. After this, they were asked to view another altered set of 20 flags – some of these old and others were new flags. These were displayed in random succession. After every image in the altered set, participants were asked if they recognised the flag from the previous set or not.

Subsequently, the second phase (trial) of the experiment commenced in which 60 pictures were shown on-screen – 20 faces, 20 devices and 20 scenes which were displayed randomly. Once again the participants were instructed to attentively watch the stimuli that appeared on screen. Again, each of these stimuli remained on the screen for three seconds. Thereupon, a break of 30 minutes began, during which all participants carried out the common task of watching a TedxTalk – The power of vulnerability (Brown, 2010). After the break, participants were asked to view an altered set of the 60 images shown before the break – some images had been shown previously, the others were new. After each stimulus, they were asked to indicate whether they had seen the picture before or not. Hereafter, the participants filled out the provided questionnaire.

As a result of the circumstances caused by the COVID-19 pandemic and the appointed precautions, the study had to be performed over internet telecommunication applications (e.g. Skype). In order to conduct the experiment using these methods, the researchers would employ screen-sharing to allow the participants to view and respond to the images. The answers given by participants would be filled in by the researcher. As both the consent form and the questionnaire were available online, they could be completed by the participant.

Data Analysis

The ability of people to correctly discriminate between old and new stimuli was calculated as a discriminability index (d’ = ‘dee-prime’). The greater the value of d’, the higher the ability to correctly discriminate old stimuli from new stimuli. These indexes were generated using the false alarm rate (the number of ‘recognised’ new stimuli) and the hit rate (correctly recognised stimuli), in accordance with Macmillan and Creelman (2004). A d’ index below 0 indicates that participants cannot distinguish old stimuli from new stimuli. A d’ ranging from 0 to 1 indicates that participants struggle to discriminate between stimuli. With a value of over 1, participants could discriminate appropriately, whilst having a value over 2 indicates good discriminatory ability (Macmillan & Creelman, 2004). Descriptive statistics of d’ were used to explore participant’s ability to discriminate between old and new stimuli, taking into consideration if the stimulus of trustworthy or not. The d’ from the pre-trial phase gave an indication of a participants’ memory ability.

Having calculated d’ and using these values, detector groups were created. The levels of d’ were used to develop three groups (equally good at detection, better trustworthy detectors and better untrustworthy detectors) for each type of stimuli and all stimuli. This was done by subtracting the d’ for the trustworthy stimuli from the d’ for untrustworthy stimuli. Consequently, values over 0 represent people who are better untrustworthy detectors. Through subtracting the d’ for untrustworthy stimuli from d’ of trustworthy stimuli indicated values below 0 as representing people who are better trustworthy detectors.

Descriptive statistics and frequencies were computed in SPSS for age, gender, the respective detector groups and the discriminability indexes (d’). Repeated measures ANOVA analyses were employed to measure a person’s ability to recognise old stimuli from new stimuli and compare it to their memory ability. This was done by using d’ of trustworthy and d’ of untrustworthy as within-subjects factors and the memory ability as a covariate. In order to analyse the level of trust in technology in the respective detector groups, one-way ANOVA analyses were also employed – the dependant variable being the level of trust in technology and the detector groups being the fixed factors. A trust variable was computed by using the mean of the answers given to all seven questions in the trust questionnaire. This was used to test if trust in technology has an influence on the participant’s ability to discriminate between old and new stimuli.

Results

Discriminability Between Old and New Stimuli

The descriptive statistics of d’ of the pre-trial phase showed that most participants on average seem to have no trouble in discriminating old and new stimuli (*M* = 1.51; *SD* = .39; *SE* = .04). The mean of d’ of all stimuli, scenes and devices were only slightly lower than that of the pre-trial phase but still above 1. Faces were the only type of stimuli where most participants struggled distinguishing between what had been shown before and what was new. Yet still, on average untrustworthy faces (*M* = .9; *SD* = .54; *SE* = .06) seem to be better recognised than trustworthy faces (*M* = .85; *SD* = .58; *SE* = .06). This finding was also apparent with all stimuli, scenes and devices – untrustworthy pictures had a higher mean of d' than the trustworthy pictures of the given category (see Table 1). These descriptive statistics indicate some sort of effect, however, further tests did not support this; no significant statistical result substantiated this.

Table 1

Descriptive statistics of dee-primes

	Pretrial	**Overall**		**Faces**		**Scenes**		**Devices**	
	d'pretrial	*d'trust*	*d'untrust*	*d'trust*	*d'untrust*	*d'trust*	*d'untrust*	*d'trust*	*d'untrust*
N	83	83	83	83	82	82	83	81	82
Mean	1.51	1.31	1.34	.85	.9	1.31	1.18	1.14	1.14
Std. Error	.04	.05	.05	.06	.6	.05	.06	.05	.05
SD	.39	.5	.5	.58	.54	.47	.56	.47	.52
Minimum	.15	0	0	-.88	-.39	-.3	-.4	0	-.47
Maximum	1.95	2.23	2.2	1.8	1.94	2.06	2.06	2.06	2.06

Ability of detection

Descriptive statistics of the detector groups represent the amount of equally good detectors of (un)trust, better untrustworthy detectors and better trustworthy detectors in the given sample. These frequencies indicate that there are on average some differences in the population. The percentages show that in the overall trial more people seem to remember trustworthy stimuli more (55.4%) than people who remembered untrustworthy more (43.4%) (see Appendix D). However, additional tests that have been conducted do not support this.

The repeated measures ANOVA analysis of all trustworthy and untrustworthy stimuli as within factor levels and memory ability as the covariate was non-significant within-subjects [$F(1, 82) = .23$; $p = .635$; $\eta_p^2 = .003$]. However, tests of between-subjects effects was significant [$F(1, 82) = 13.93$; $p < .001$; $\eta_p^2 = .147$]. Similar results were found when running an ANOVA analysing the scene stimuli - it was non-significant within-subjects [$F(1, 81) = .28$; $p = .599$; $\eta_p^2 = .003$]. Yet, between-subjects of scene stimuli and memory ability was significant [$F(1, 81) = 12.502$; $p = .001$; $\eta_p^2 = .135$]. This means that the overall values differ from 0, therefore, people are able to memorise. People are able to recognise old from new, which is also confirmed by the control stimuli in the pre-trial. Thus, memory did assist in detection, though there is no substantial difference in the correct recognition of trustworthy and untrustworthy pictures – participants were equally good at both.

The repeated measures ANOVA analyses was non-significant for both device stimuli within-subjects [$F(1, 80) = .47$; $p = .496$; $\eta_p{}^2 = .006$] and face stimuli within-subjects [$F(1, 81) = .87$; $p = .353$; $\eta_p{}^2 = .011$]. Moreover, there is no significant difference in detection between the untrustworthy and trustworthy faces [$F(1, 81) = 1.93$; $p = .168$; $\eta_p{}^2 = .024$]. No significant difference in detection of untrustworthy or trustworthy devices was found [$F(1,80)=3.08$; $p=.083$; $\eta_p{}^2 =.038$]. Based on these results, when viewing faces and devices memory did not assist in detection and no difference in the ability to recall trustworthy and untrustworthy stimuli.

Trust and Detection Groups

The predisposition to trust in technology, measured by the trust questionnaire, and the tendency of groups of trustworthy and untrustworthy discriminators was also investigated. Doing so will firstly shed light on whether future research on trust before the use of technology should incorporate this variable, and secondly if this variable could explain the difference in people that have a tendency to recall (un)trust better in the different types of stimuli.

General linear models of general trust in technology and the ability to detect were all non-significant, yet some tendencies seem to be evident. Within the detection groups of overall stimuli, face stimuli, and device stimuli, it seems that people that are better at remembering untrustworthy stimuli have a lower score on the trust scale and people that are better at remembering trustworthy stimuli to have a higher score on the trust questionnaire. Contrastingly, the resulting model of trust and scene detection, which was also non-significant, showed that people that are better at remembering untrustworthy scenes seem to have a higher score on the trust scale than people who are better at remembering trustworthy scenes.

Discussion

The present study's goals were to explore whether previously categorised (un)trustworthy stimuli are recognised better and whether the trustworthiness of a stimulus influences memory. In addition to this, it was explored if the inclination to recognise trustworthy or untrustworthy stimuli more is influenced by a person's level of trust in technology. The present study also analysed if the type of stimuli (faces, scenes, devices) is an important factor in regards to (un)trust and recognition of old and new stimuli. The results, limitations and practical implications of this exploratory study are discussed in more detail below.

The results of the present study, although non-significant, are noteworthy. At first sight, descriptive statistics of the discriminability indexes seem to indicate that on average untrustworthy stimuli are more easily recognised than trustworthy ones. This is in line with research that discusses the survival-mechanisms of humans to have the predisposition to remember untrustworthy events and people more (McBride et al., 2013; Hou, & Liu, 2019; Kroneisen, 2018). Untrustworthy faces are imprinted in our long-term memory, and viewing these faces even influences people's encoding and retrieval (Weymar et al., 2019). However, further tests in this study could not verify this phenomenon. Nevertheless, the trustworthiness of a given stimulus appears to have some sort of impact on people's decision-making.

Furthermore, the memory ability measured in the pre-trial seems to have some sort of effect on the ability to distinguish between old and new stimuli. This seems to also be the case when looking at scenes. This is in line with the literature found on memory and trust. Memory has a certain significance in recognition of (un)trustworthiness. Previous literature states that being able to recognise cheaters is more important than being able to recognise cooperators (non-cheaters) (Verplaetse et al., 2007). Similarly, the ability to recognise what is not worthy of trust may be more beneficial for people than being able to recognise what is trustworthy. However, the outcome of these analyses between detector groups and the trust scale score of participants display some interesting effects. A general sense of trust in technology does not affect people's ability to discriminate what is trustworthy or not. This poses the possibility of the mentioned cheater detection mechanism being different from the concept of untrust.

Even though studies have confirmed that trust in technology has an immense effect on how people choose to interact with technology and that it is imperative to the adoption of technology (Muir & Muray, 1996), the present study has not been able to substantiate this using the McKnight trust questionnaire. This finding is very relevant for understanding and furthering TTS and research in this field. In the study of Muir and Muray (1996), results showed that user trust in a device was mostly based on the perceived competence of the machine. That no effect was found when investigating if trust does influence recognition of (un)trustworthy stimuli, could be explained by the trust questionnaire used in the present study. The questionnaire does not have any items explicitly pertaining to perceived competence of a system. Additionally, the trust scale used in the present study was specifically designed to measure trust in technology and, therefore, only takes items regarding functionality, reliability and helpfulness into account. None of these factors seems to be correlated to a person's level of trust in a given environment.

Accordingly, the finding that trust in technology does not influence detection ability of either old and new scenes and faces may be explained by the fact that the items are of no significance when viewing these two types of stimuli.

Limitations and future directions

The present exploratory study does have three limitations. Due to the COVID-19 pandemic, the procedure of this study had to be adapted to the circumstances. Therefore, the experiment had to be performed with the participant over video chat using screen sharing. This was not optimal, as there is no way of knowing if the participants had been able to view the stimuli for the specified time or if internet lags had interfered. Moreover, between the presentation of stimuli and the recall of stimuli, there was a break of only 30 minutes. Redesigning the experiment in order to better analyse whether there is a substantial difference in the correct recognition of old stimuli amongst new, extending this break to perhaps a couple of weeks may give more insight. In this way, long-term memory and its impact on the detection ability of (un)trust could be explored. In addition to this, in the present experiment participants are asked whether they have seen a certain stimulus before. A more potent measure of untrust, in this case, would be to add another variable asking whether they think the given stimulus is trustworthy or not.

Future research on the subject of untrustworthiness is greatly needed, as this fairly new term has scarcely been investigated. Untrust has only in recent years been taken into account in research. To truly evaluate and understand trust, all aspects must be explored. In line with this reasoning, when replicating this experiment, it should perhaps take another trust scale into account other than the McKnight Technology Scale. The questionnaire is fairly short, and items pertain only to perceived helpfulness, functionality, and reliability of technology. These are all factors taken into account, especially when dealing with systems with less human features. However, the more humanness the technology displays, the more the assessment of trust is based on competence and integrity (Lankton et al., 2015; Muir & Muray, 1996). The forms of technology are vast. Hence, future research may benefit from adding different items that assess these concepts as well, as this may evoke very different results.

Furthermore, future work should investigate the concept of (un)trust from an evolutionary and neuroscientific point of view. The ability to remember cheaters more could be linked to an evolutionary mechanism; for the sake of survival, humans are more inclined to remember deceivers (McBride et al., 2013; Hou, & Liu, 2019). Moreover, a functional neuroimaging study by Dimoka (2010), found that distrust and trust activate two separate areas of the brain. While trust is found to be associated with the area of the brain responsible for uncertainty, reward and prediction, distrust activated areas responsible for emotions and fear of loss. In line with this, it could be hypothesised that higher activation in one area could lead to better recognition of that type of stimuli. This may explain the difference between the respective trustworthy and untrustworthy detector groups.

Lastly, a more focussed investigation of how the effect of memory on trust recognition works could be of great significance. Previous research shows the significant effect of untrustworthiness on memory and encoding, yet fails to explain how this works (Mattarozzi et al., 2015). A possible explanation for this phenomena could be, that untrustworthy stimuli may attract more attention; therefore, the encoding process for these stimuli is better, and consequently, recognition ability may be enhanced for these stimuli. This could be tested by having participants view (un)trustworthy pictures whilst doing an fMRI. In this way, the neurological and specific cognitive mechanisms behind encoding and retrieval would shed light on how (un)trust is formed and affects people's everyday life.

Conclusion

This exploratory study has contributed to research by i.) investigating whether (un)trustworthiness influences the recognition of stimuli, ii.) if these differences occur due to the type of stimuli, and iii.) examining whether the level of trust in technology influences the differences in people’s inclination to better detect either trustworthy or untrustworthy stimuli. The findings of previous research that untrustworthy stimuli are recognised more than trustworthy stimuli could not be replicated. However, the present study makes clear that the trust scale used here may not be an adequate tool to measure (un)trust. Consequently, it should be adapted in future research. In redesigning this experiment, whereby acknowledging the limitations mentioned, may yield more significant results - the level of trust a user places in technology before using can be measured. Although the topic of trust is a complex one, researching the influence of trust in technology is essential in order to enhance trust toward systems in today’s technology dependant world.

References

Alesina, A., & La Ferrara, E. (2002). Who trusts others?. *Journal of public economics*, *85*(2), 207-234.

Baba, M. L. (1999). Dangerous liaisons: Trust, distrust, and information technology in American work organisations. *Human organisation*, 331-346.

Bar, M. (2007). The proactive brain: using analogies and associations to generate predictions. Trends in Cognitive Sciences, 11(7), 280-289. doi:10.1016/j.tics.2007.05.005

Borsci, S., Buckle, P., Walne, S., & Salanitri, D. (2018). Trust and Human Factors in the Design of Healthcare Technology. Paper presented at the Proceedings of the 20th Congress of the International Ergonomics Association (IEA 2018), Cham.

Brown, B. (2010, June). The power of vulnerability. Retrieved June 29, 2020, from https://www.ted.com/talks/brene_brown_the_power_of_vulnerability?language=en

Chiappe, D., & Brown, A. (2004). Cheaters are looked at longer and remembered better than cooperators in social exchange situations. Evolutionary Psychology, 2, 108120.

Cosmides, L. (1989). The logic of social exchange: Has natural selection shaped how humans reason? Studies in the Wason Selection Task. Cognition, 31, 187276.

Cosmides, L., & Tooby, J. (1992). Cognitive adaptations for social exchange. In J. H. Barkow, L. Cosmides, & J. Tooby, (Eds.), The adapted mind: Evolutionary psychology and the generation of culture(pp. 163–228) New York, NY: Oxford University Press.

CPSC.gov. (n.d.). Retrieved June 29, 2020, from https://www.cpsc.gov/.

Crone, D. L., Bode, S., Murawski, C., & Laham, S. M. (2018). The Socio-Moral Image Database (SMID): A novel stimulus set for the study of social, moral and affective processes. *PloS one*, *13*(1), e0190954.

Davis, F.D., (1989). Perceived Usefulness, Perceived Ease of Use, and User Acceptance of Information Technology, MIS Quarterly, *13*(3), pp.319-339.

Dimoka, A. (2010). What does the brain tell us about trust and distrust? Evidence from a functional neuroimaging study. *Mis Quarterly*, 373-396.

FeldmanHall, O., Dunsmoor, J. E., Tompary, A., Hunter, L. E., Todorov, A., & Phelps, E. A. (2018). Stimulus generalisation as a mechanism for learning to trust. *Proceedings of the National Academy of Sciences*, *115*(7), E1690-E1697.

Frederiksen, M. (2011). Distrust and Mistrust: Lessons from a Low Trust Minority. Comparative Sociology, 10(2), 248-280. Doi:10.1163/156913311X56658

Friedman, B., Peter, H., Khan, J., & Howe, D. C. (2000). Trust online. *Commun. ACM, 43*(12), 34-40. doi:10.1145/355112.355120

Hosmer, L. T. (1995). Trust: The connecting link between organisational theory and philosophical ethics. *Academy of management Review*, *20*(2), 379-403.

Hou, C., & Liu, Z. (2019). The survival processing advantage of face: The memorization of the (un) trustworthy face contributes more to survival adaptation. *Evolutionary Psychology*, *17*(2), 1474704919839726.

Knack, S., & Keefer, P. (1997). Does social capital have an economic payoff? A cross-country investigation. *The Quarterly journal of economics*, *112*(4), 1251-1288.

Kramer, R. M. (1999). Trust and distrust in organizations: Emerging perspectives, enduring questions. *Annual review of psychology*, *50*(1), 569-598.

Kroneisen, M. (2018). Is he important to me? Source memory advantage for personally relevant cheaters. *Psychonomic bulletin & review*, *25*(3), 1129-1137.

Lankton, N., McKnight, D., & Tripp, J. (2015). *Technology, Humanness, and Trust: Rethinking Trust in Technology* (Vol. 16).

Lewicki, R. J., McAllister, D. J., & Bies, R. J. (1998). Trust and distrust: New relationships and realities. *Academy of management Review*, *23*(3), 438-458.

Luhmann, N. (1979). Trust and Power New York Wiley.

Ma, D., Correll, J. & Wittenbrink, B. (2015). The Chicago Face Database: A Free Stimulus Set of Faces and Norming Data. Behavior Research Methods, 47, 1122-1135.

Macmillan, N. A., Creelman, C. D. (2004). Detection theory: A user's guide. Psychology press.

Marsh, S., & Dibben, M. R. (2005, May). Trust, untrust, distrust and mistrust–an exploration of the dark (er) side. In *International conference on trust management* (pp. 17-33). Springer, Berlin, Heidelberg.

Mattarozzi, K., Todorov, A., Codispoti, M. (2015). Memory for faces: the effect of facial appearance and the context in which the face is encountered. *Psychological Research :An International Journal of Perception, Attention, Memory, and Action*, 79(2), 308-317.doi:10.1007/s00426-014-0554-8

McBride, D. M., Thomas, B. J., & Zimmerman, C. (2013). A test of the survival processing advantage in implicit and explicit memory tests. *Memory & cognition*, *41*(6), 862-871.

McKnight, D. H. (2005). Trust in information technology. *The Blackwell encyclopedia of management*, *7*, 329-331.

McKnight, D. H., Choudhury, V., & Kacmar, C. (2002). Developing and validating trust measures for e-commerce: An integrative typology. *Information systems research*, *13*(3), 334-359.

McKnight, D. H., Carter, M., Thatcher, J. B., & Clay, P. F. (2011). Trust in a specific technology: An investigation of its components and measures. *ACM Transactions on management information systems (TMIS)*, *2*(2), 1-25.

Mealy, L., Daood, C., & Krage, M. (1996). Enhanced memory for faces of cheaters. Ethology and Sociobiology,17, 119128.

Muir, B.M. and Moray, N., 1996. Trust in automation. Part II. Experimental studies of trust and human intervention in a process control simulation. Ergonomics, 39(3), pp.429 460.

Oda, R. (1997). Biased face recognition in the prisoner's dilemma games. Evolution and Human Behavior,18, 309317.

Rousseau, D. M., Sitkin, S. B., Burt, R. S., & Camerer, C. (1998). Not so different after all: A cross-discipline view of trust. *Academy of management review*, *23*(3), 393-404.

Rule, N. O., Slepian, M. L., & Ambady, N. (2012). A memory advantage for untrustworthy faces. *Cognition*, *125*(2), 207-218.

Sasse, M. A. (2005). Usability and trust in information systems. Edward Elgar.

Vance, A., Elie-Dit-Cosaque, C., & Straub, D. W. (2008). Examining trust in information technology artifacts: the effects of system quality and culture. *Journal of management information systems*, *24*(4), 73-100.

Verplaetse, J., Vanneste, S., & Braeckman, J. (2007). You can judge a book by its cover: the sequel.: A kernel of truth in predictive cheating detection. *Evolution and Human Behavior*, *28*(4), 260-271.

Wang, W., & Benbasat, I. (2005). Trust in and adoption of online recommendation agents. *Journal of the Association for Information Systems*, *6*(3), 72-101.

Wendt, J., Weymar, M., Junge, M., Hamm, A. O., & Lischke, A. (2019). Heartfelt memories: Cardiac vagal tone correlates with increased memory for untrustworthy faces. *Emotion*, *19*(1), 178.

Weymar, M., Ventura-Bort, C., Wendt, J., & Lischke, A. (2019). Behavioral and neural evidence of enhanced long-term memory for untrustworthy faces. *Scientific Reports*, *9*(1), 1-8.

Appendix A

Demographics

1. What is your age?

2. What is your sex?
    a. Male
    b. Female
    c. Other

3. What is your nationality? (e.g. Dutch, English, German etc.)

4. What is the highest level of school you have completed or the highest degree you have received?

a. Less than high school degree
b. High school graduate (high school diploma or equivalent including GED)
c. Some college but no degree
d. Associate degree in college (2-year)
e. Bachelor’s degree in college (4-year)
f. Master’s degree
g. Doctoral degree
h. Professional degree (JD, MD)

5. What is your occupation/ What do you study?

Appendix B

General Trust in Technology Measurement

Faith in General Technology (Adapted from McKnight et al. 2002):

1. I believe that most technologies are effective at what they are designed to do.
2. A large majority of technologies are excellent.
3. Most technologies have the features needed for their domain.
4. I think most technologies enable me to do what I need to do.

Trusting Stance—General Technology (Adapted from McKnight et al. 2002):

1. My typical approach is to trust new technologies until they prove to me that I shouldn't trust them.
2. I usually trust a technology until it gives me a reason not to trust it.
3. I generally give a technology the benefit of the doubt when I first use it

Appendix C

Examples of Stimuli Used in the Experiment

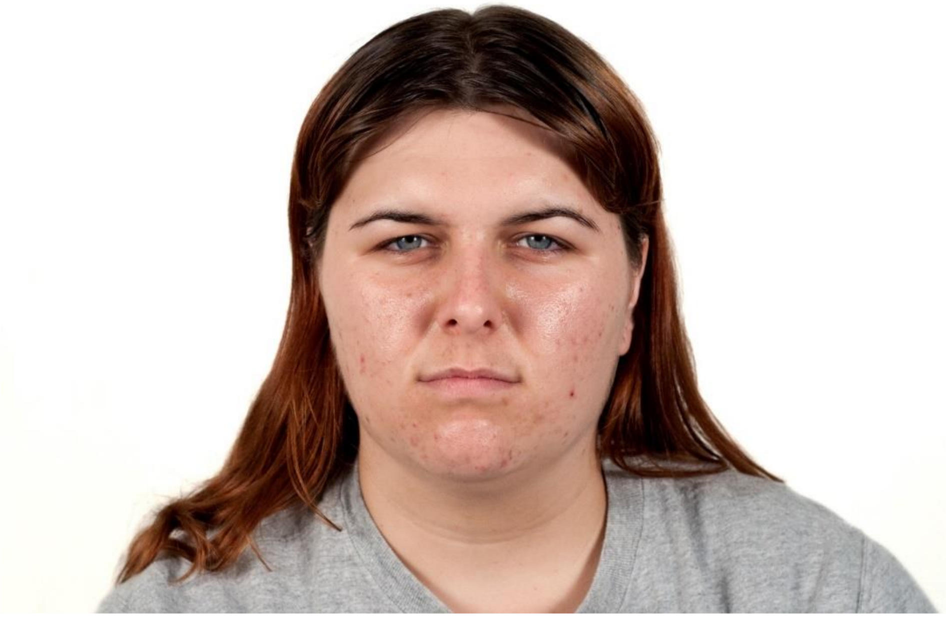

Figure 1. Example of an untrustworthy face used in the experiment.

Figure 2. Example of a trustworthy face used in the experiment.

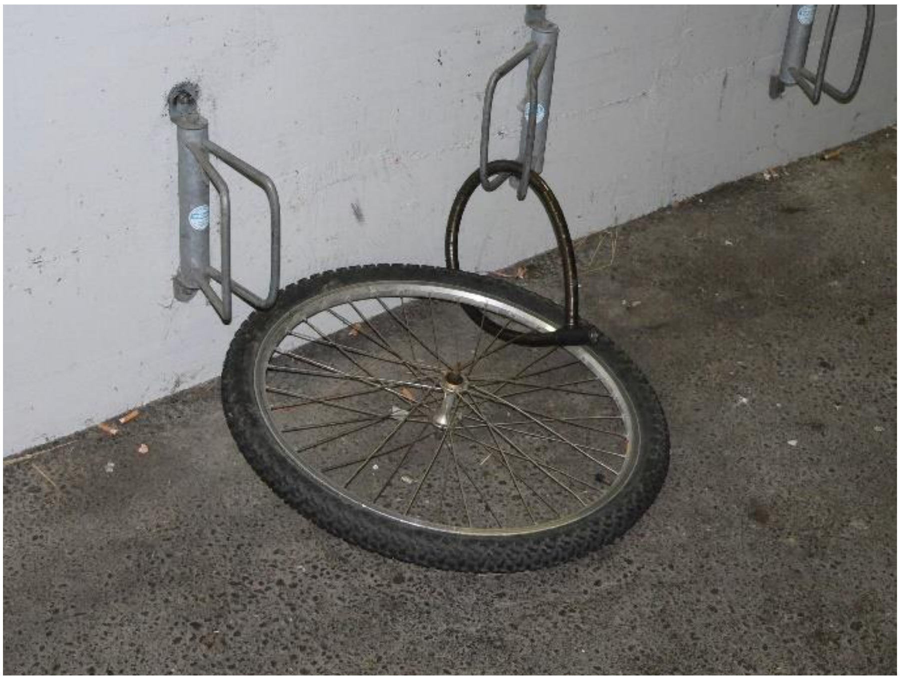

Figure 3. Example of an untrustworthy scene used in the experiment.

Figure 4. Example of a trustworthy scene used in the experiment.

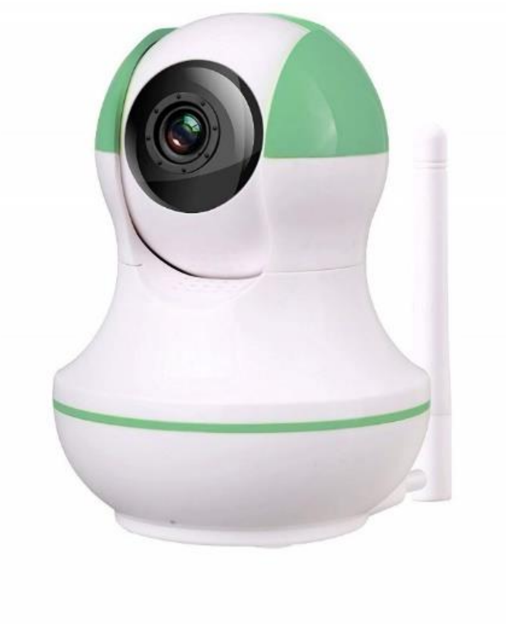

Figure 5. Example of an untrustworthy device used in the experiment.

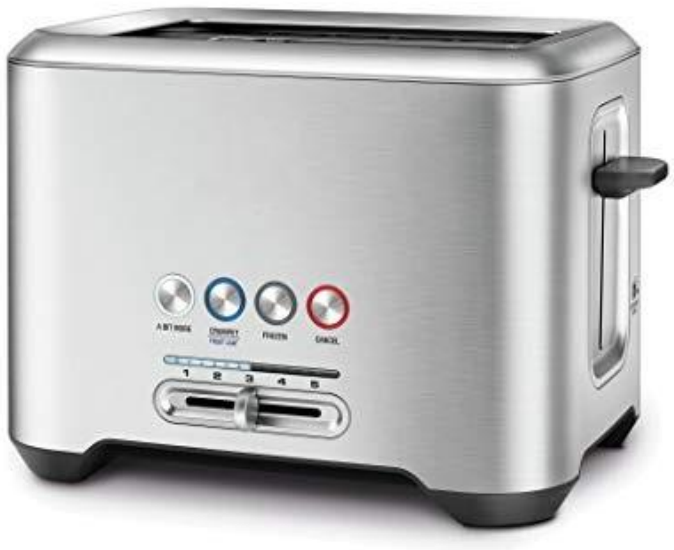

Figure 6. Example of a trustworthy device used in the experiment.

Appendix D

Descriptive Statistics

Table 2

Descriptive statistics of overall detectors

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	0.00	1	1.2	1.2	1.2
	1.00	36	43.4	43.4	44.6
	2.00	46	55.4	55.4	100.0
	Total	83	100.0	100.0	

Note. Groups respectively being 0 = equally good at detecting trustworthy and untrustworthy stimuli, 1 = better detection of untrustworthy stimuli, 2 = better detection ability of trustworthy stimuli

Table 3

Descriptive statistics of scene detectors

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	0.00	8	9.6	9.8	9.8
	1.00	33	39.8	40.2	50.0
	2.00	41	49.4	50.0	100
	Total	82	98.8	100.0	
Missing	System	1	1.2		
Total		83	100.0		

Note. Groups respectively being 0 = equally good at detecting trustworthy and untrustworthy stimuli, 1 = better detection of untrustworthy stimuli, 2 = better detection ability of trustworthy stimuli

Table 4

Descriptive statistics of face detectors

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	0.00	1	1.2	1.2	1.2
	1.00	45	54.2	54.9	56.1
	2.00	36	43.4	43.9	100.0
	Total	82	98.8	100.0	
Missing	System	1	1.2		
Total		83	100.0		

Note. Groups respectively being 0 = equally good at detecting trustworthy and untrustworthy stimuli, 1 = better detection of untrustworthy stimuli, 2 = better detection ability of trustworthy stimuli

Table 5

Descriptive statistics of device detectors

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	0.00	3	3.6	3.7	3.7
	1.00	41	49.4	50.6	54.3
	2.00	37	44.6	45.7	100.0
	Total	81	97.6	100.0	
Missing	System	2	2.4		
Total		83	100.0		

Note. Groups respectively being 0 = equally good at detecting trustworthy and untrustworthy stimuli, 1 = better detection of untrustworthy stimuli, 2 = better detection ability of trustworthy stimuli