MASTER THESIS

An Exploratory Study on Recognition of Untrustworthy Devices

Jaume Agud Morera

Human Factors and Engineering Psychology
Faculty of Behavioral Management and Social Sciences

Tutors
Dr. S. Borsci
Dr. R.H.J. van der Lubbe

Enschede, 25-01-2021

UNIVERSITY OF TWENTE.

Abstract

People seem to rely on appearance as an indicator of trustworthiness. Previous research showed that people can remember images of untrustworthy people better than images of trustworthy people, suggesting people's capacity to recognize what not to trust. Whether the same effect occurs with other visual stimuli, such as images of devices or scenes, is unexplored. This research aimed to explore whether there is a difference in remembering trustworthy or untrustworthy stimuli after being exposed to pictures of faces, devices, and scenes. Through a memory experiment, both the differences in memory as well as the underlying individual factors affecting memory performance were explored. The approach of Verplaetse et al. (2007), who studied the memory advantage for faces of cheaters using a memory task, was applied, expecting an enhanced memory towards pictures of untrustworthy stimuli. A memory test was carried out in which a total of sixty images (mixing faces, scenes, and devices, both trustworthy and untrustworthy) were shown twice, with a thirty-minute break, to a total of thirty participants. Each image was previously classified as trustworthy or untrustworthy, and participants were unaware of this categorization. The results showed no differences in remembering trustworthy or untrustworthy stimuli. Also, none of the personality factors analysed (Technology Acceptance, Geekism and Trust Score) was correlated to memory recognition. This outcome suggests that people cannot predict how trustworthy a device is based on appearance and adds some recommendations for future research.

Keywords: Cheater detection mechanism, Trust Towards Systems, Memory recognition, Appearance.

Contents

Introduction

What do personal relations, business transactions, and complex social organizations have in common? These are not only some of the most challenging activities that humankind must deal with, but also, examples of some of the spheres of life in which trust is a paramount and critical factor (Hosmer, 1995). Given its crucial role in human life, trust has been thoroughly researched for decades, from studies on its foundational factors (Mayer et al., 1995), to the attempts of deciphering how people decide if and who to trust in a particular situation (Das & Teng, 2004).

Trust can be defined as "the willingness of a party to be vulnerable to the actions of another party, based on the expectation that the other will perform a particular action" (Mayer et al., 1995, p. 715). Although trust has traditionally been associated with interactions amongst humans, researchers such as McKnight et al. (2011) or Thatcher et al. (2011) have argued that humans can also develop a sense of trust towards systems, which is defined as a set of beliefs before experiencing a system, built through relationship and dependent on experience (Borsci et al., 2018).

Trust Towards Systems is built through experience, however, it exists and starts developing already before people interact with a system (Borsci et al., 2018). This concept is referred to as "*Pre-use Trust Towards Systems*" and includes a set of expectations that develop via an indirect exposure to a product, such as advertising, word of mouth or previous experiences (Borsci et al., 2018). These factors influence how technology is perceived prior to an interaction (McKnight et al., 2011; Salanitri et al., 2015). When these cues are missing, people might make use of other available cues linked to the product trustworthiness, such as its appearance (Borsci et al., 2018), influencing the perceived trustworthiness of the product (Pengnate & Sarathy, 2007) and biasing expectations.

So, people seem to make judgements on the trustworthiness of technology based on appearance, which can be risky as these assessments might be proven wrong and change dramatically after a device has been used. This is often caused by a negative experience, in which the device failed to perform as expected or caused damage to the user (Borsci et al., 2018). A better understanding of people's ability to detect untrustworthy technology would help in preventing these risks.

The present study aims to explore whether there is a difference in remembering untrustworthy and trustworthy stimuli, as well as the individual factors affecting recognition of trustworthy and untrustworthy stimuli. Although the study of trust towards technology before the usage is an emerging topic which has been minimally investigated, and it is still unclear how it is developed, some studies have shown that people can recognize what not to trust. For instance, experiments using a memory test with pictures of faces of cheaters or co-operators showed that participants had an enhanced memory towards pictures of faces of cheaters, in comparison to pictures of faces of co-operators.

According to Verplaetse et al. (2007), this evidences that people can recognize cheaters, aided by predictive cognitive modules that work as an automatic processing skill. Thus, certain cognitive modules are adapted to deal with cheaters, defined as "individuals who intentionally violate a social contract by taking the benefit specified without satisfying the requirement that provision of that benefit was made contingent on" (p. 200), to avoid them in future transactions.

In the case of technology, a cheater is an unreliable device which does not perform as expected (Nickel et al., 2010) and poses a risk to the final user (Borsci et al., 2018). The underlying assumption of this research is that, if people can assess how trustworthy another person is based on its facial appearance (Todorov et al., 2009), a similar predictive mechanism might help people in selecting and judging a new technology on basis of its appearance. This would lead to an enhanced memory towards the untrustworthy devices, compared to the trustworthy devices, similarly to what has been observed in memory experiments with faces of cheaters and faces of cooperators.

Human Trust and Trust Towards Technology

Given the absence of previous literature addressing people's ability to recognize untrustworthy devices based on visual cues, this research aims to apply the methods and theories utilized in the studies on cheater detection in traditional trust. Although trust towards systems and traditional trust are different in some critical points (Lippert & Swiercz, 2007), its similarities justify the possibility to apply the same methods in the study of untrustworthiness detection.

The most evident difference between traditional trust and trust towards systems is the object of trust, which in the case of TTS is a device or system, instead of another person (Lee & Turban, 2001; McKnight et al., 2011). This means that trust exists in one direction, as technology cannot trust in return, and that this trust is based only on non-volitional and amoral factors, because technologies cannot decide if they want to cooperate or defect purposely. According to Thatcher et al. (2011), Trust Towards Systems is driven by three "object-related" beliefs:

- Functionality belief: The system has the capability and features to do what needs to be done.
- Helpfulness belief: It is the belief that the system will provide adequate and responsive aid.
- Reliability belief: Refers to the belief that a system will act in a consistent and predictable way.

All these beliefs are, respectively, equivalent to the concepts of competence, benevolence, and integrity, used to measure trust in humans (McKnight et al., 2011). This showcases that, besides its differences, Trust Towards Systems and Trust Towards Humans are both grounded on the same beliefs and respond to identical psychological needs. In fact, some studies draw an existing connection, reporting that a higher trust in people is correlated with a higher TTS (Gefen, 2000; Teo & Liu, 2007). This correlation also suggests that the general cognitive mechanisms responsible for an enhanced memory might be useful not only for avoiding cheaters in a situation of social exchange, but a wider range of other harmful stimuli, as Bell & Buchner (2012) proposed.

According to different authors (Verplaetse et al., 2007; Yamagishi, 2003), one of the cognitive mechanisms that humans are equipped with is responsible for providing an enhanced memory for faces or cheaters, to avoid them in the future. As explained before, from an evolutionary point of view, it could be possible that this mechanism has evolved, allowing people to discriminate (and remember better) a wider range of potentially dangerous stimuli (Bell & Buchner, 2009). If evolution works to improve the efficiency of cognitive mechanisms, it could be expected that humans have developed the skill to identify untrustworthy devices. In order to study the possible memory advantage for untrustworthy devices, and in absence of other methodologies available to this end, it seems possible to apply the same methods that are used to study memory advantage for faces of cheaters. Two main arguments justify why this is possible.

First, in both cases the question to be addressed is the same: Whether people can identify cheaters based on limited available information. Just like in some cases people decide to trust another person based on limited information (Oosterhof &Todorov, 2009), people also must assess occasionally the trustworthiness of a device prior to interacting with it (Borsci et al., 2018). Different factors play a role in these cases: Some of them are inherent to the trustor, such as predisposition to trust or previous experiences (McKnight et al., 2011), whilst others depend on the trustee. This happens both when the object of trust is a device or another person.

Second, people only have access to visual cues. If in traditional trust, facial appearance helps people to make a judgement about how trustworthy another person is (Oosterhof & Todorov, 2009), in trust towards systems is the design of a device which provides valuable information for people to shape their expectations regarding the trustworthiness, usability and performance of a device

(McKnight et al. 2011; McKnight et al. 2002; Lankton et al., 2015; Salanitri et al., 2015). For example, Fogg et al. (2002) found that the credibility of a website was partly based on the appeal of the overall visual design. Similarly, Harley (2016) found that design quality positively affects the perceived trustworthiness in web design. From a more general perspective, Tractinsky et al. (2000) proposed the existence of a halo effect of aesthetics, which means that physical beauty tends to affect later perceptions and inferences about other traits. Therefore, appearance would affect the perceived trustworthiness of a technology prior its usage.

The Role of Memory in Trust related studies

Several methods can be employed to study if people are able to detect untrustworthy counterparties. As Herse et al. (2018) stated, investigating trust is complicated, especially when it comes to studying Trust Towards Systems (TTS). One way of approaching this task is by utilising indirect measures to investigate trust, by which trust is assessed using a disguised method of non-obtrusive behavioural observation (Herse et al., 2018), instead of utilizing direct questionnaire. According to Glaeser et al. (2000) disguised methods can be more effective and resistant to bias than direct methods. The usage of indirect measures to study trustworthiness detection has been a common practice in the field of social psychology, with memory standing as one of the preferred tools used by different researchers (Verplaetse et al., 2007).

But why and how can memory indicate if people can detect potentially dangerous stimuli? The first approximation to this question is grounded on the theories from Cosmides (1989) and Cosmides and Tooby (1992), who proposed the existence of a cognitive mechanism that would allow people to remember (better) those who violated social contracts, in order to avoid them in the future. From an evolutionary point of view, one of the most important tasks humankind has had to deal with historically was to discriminate between cheaters and co-operators (Mealey et al., 1996). Consequently, evolution might have designed the human mind to scan for information that might signal intentions to defect (in other words, untrustworthy counterparties). This would result in increased attention towards no cooperativeness signals and, therefore, an increased memory towards untrustworthy counterparties (Verplaetse et al., 2007).

In line with the previously presented theories of Comides (1989) and Cosmides & Tooby (1992), Mealey et al. (1996) performed a pioneer study to test the hypothesis of enhanced memory for faces of cheaters. They found that participants were able to remember better pictures of faces associated with a story of cheating (untrustworthy) than faces of subjects with a story of trustworthiness, after presenting them a week later. This supported the assumption of the existence of selective attention and storage mechanisms for processing social information. With similar goals and methods, Oda (1997) tested the memory advantage for threatening faces in a context of cooperation. They found that male cheaters who participated in a Prisoner's dilemma game were remembered better than male co-operators with a robust effect of a biased face recognition towards cheaters.

Both the experiments of Mealey et al. (1996) and Oda (1997) provided participants with explicit information about the degree of trustworthiness of the subjects (descriptions and stories). Using a different approach, Yamagishi et al. (2003) argued that people would be able to remember defectors' faces better than those of co-operators without being told who defectors or co-operators are since they look different and people are able to detect this. After four experiments, they concluded that subjects recognized faces of defectors better than those of co-operators, even though they did not know which faces were of defectors. This suggested that some facial features distinguish defectors from co-operators and that people can consciously identify such features. This would be in line with the postulations of Davey (2005), who argued the existence of general, non-conscious mechanisms evolved to facilitate attention to the thread-related stimulus.

Extending the scope of the previous studies, the experiments of Verplaeste et al. (2007) showed that people are better at remembering faces of cheaters than co-operators and that this evidences people's ability to detect cheaters better than co-operators. These results could be interpreted as an indication of either the existence of a cheater detection mechanisms (Mealey et al., 1996), or to general mechanisms that favour the information with greater diagnostic value (Chiappe et al., 2004; Bell & Buchner, 2009).

Aim of the Present Study

The present exploratory research intents to explore whether people also have an enhanced memory for untrustworthy devices compared to trustworthy devices. In line with previous research, it would be expected that people can remember better images of untrustworthy devices than images of trustworthy devices. Building upon Verplaetse et al. (2007), a memory test was developed and tested using as stimuli images of faces, scenes, and devices. These images were subcategorised prior to the test as either trustworthy or untrustworthy. In line with previous studies on cheater detection (Yamagishi et al., 2003; Verplaetse et al., 2007), several measures of discrimination(d') were calculated for each participant to assess the ability to recognize pictures of different stimuli as old or new.

Two questions drove our exploratory analysis. The first question can be summarised as follows: Can people remember the images of untrustworthy devices better, as compared to images of trustworthy devices?

Following the previous studies, if the memory mechanism responsible of the memory advantage for cheaters acts by highlighting relevant information, it might also be useful with other potential harmful stimuli (Bell & Buchner, 2009). Then, a higher memory for untrustworthy stimuli is expected, similarly to what has been obtained with faces (Mealey et al., 1996; Verplaetse et al., 2007).

In addition to identify whether some participants can remember better images of untrustworthy stimuli, this study also aimed to explore the personality factors that might cause this effect in memory. Therefore, and grounded on the idea that TTS is affected by some personality traits, like self-esteem, self-efficacy, and capabilities with similar technologies (Borsci et al., 2018), this research explored the effect of 3 personality factors on memory performance:

- The first factor is the Trust Score. Some studies showed that the propensity to trust has a positive effect on online trust formation (Gefen, 2000; Teo & Liu, 2007).
- Secondly, Technology acceptance, which is influenced by the trust (Salam et al., 2005), and equally important when adopting new technology (Gefen et al., 2003).
- The third factor is Geekism. Metzger et al. (2013) and De Angeli et al. (2006) pointed out that proficiency, experience, and expertise with products can affect the trustworthiness assessment positively.

The second question can be therefore defined as follows: Are there any individual differences that explain why some people might be better at remembering untrustworthy stimuli, compared to trustworthy stimuli?

Methods

Participants

The study sample consisted of 30 people (Females:11, Age mean: 29.9 years, SD:11.6). The sample was recruited using stratified sampling and aimed to include participants with different degrees of technology affinity, to study differences in performance based on personality factors.

Materials

Stimuli

A total of three types of stimuli were used: faces, products, and scenes. These stimuli were categorised as trustworthy or untrustworthy. The categorization was based on the available information, which differed per type of stimuli in both the criteria and source used. Below, a description and details of the categorization per type of stimuli can be found.

Faces*.* The Images of faces were retrieved from the Chicago Face Database (Ma et al., 2015), a free database consisting of pictures of faces with a neutral expression. The CFD rated the total of 598 faces per perceived trustworthiness, and the twenty highest and lowest rated faces on their scale were picked for this study, adhering to the original rates. Including images of faces permitted to replicate previous studies (Yamagishi et al., 2003; Verplaetse et al., 2007), on basis of which it was explored whether the memory enhancement for images of cheaters would also happen in the case of the two new set of stimuli (Devices and Scenes).

Scenes*.* The 40 pictures of scenes were retrieved from the Socio-Moral Image Database (SMID) (Crone et al., 2018), choosing 20 trustworthy and 20 untrustworthy scenes. These stimuli were quite diverse, portraying a mix of objects, landscapes, and people. To classify pictures as trustworthy or untrustworthy, the moral foundations theory (MFT) was used as a framework. As Crone et al. (2018) explain, there are 5 innate moral values. One of them is fairness, which concerns the identification of cheating and exploitation. Based in this, the images of the SMID rated as low in fairness were untrustworthy, and the images rated as high in fairness were trustworthy. This approach was adopted to study the memory differences for images of untrustworthy scenes, although it is not a common practice.

Devices*.* The images of devices were mostly retrieved from the Consumer Product Safety Commission (CPSC), which has a given authority by the government to recall and label untrustworthy products that pose a risk for consumers. A total of 15 different devices were retrieved from the list of recalled devices, namely, those that caused any form of damage or posed a risk to customers. After this, a trustworthy device of the same category (e.g., coffee machines) was chosen using online retail websites. 5 devices were included from other sources (U.K. government, Australian product safety commission, independent press or company recall. The inclusion criteria for the untrustworthy devices consisted of: (1) Products recalled by the CPSC (Mostly), U.K. government, Australian product safety commission, press release or independent press as being dangerous due to the occurrence of: critical failures, caused injures or other hazards. (2) Products created within the last ten years, to ensure that the perceived untrustworthiness was not due to an old/outdated design. (3) Products recalled in the last five years. Conversely, an exclusion criterion for the untrustworthy devices was developed, following these points: (1) Products whose problems could be solved by repairing or changing one of their parts (i.e., some cars). (2) Products recalled by less than ten people or with no real evidence of problems occurred. (3) Products whose design looked notably old/outdated. The trustworthy devices were selected in accordance with three criteria: Availability in online retail websites, listed for at least

1 year, and with consistent and positive reviews. These criteria were decided altogether with for experts of human factors, selecting consensually the final list of trustworthy and untrustworthy devices.

Pre-Test Material

A demographic questionnaire was developed using the Qualtrics online software system. Here, basic demographic data from the participants (age, gender, level of education, affinity with technology) was collected. In addition, the questionnaire incorporated reduced versions of three scales of personality (Trust Score, Technology Acceptance and Geekism). These were filled in using a Likert scale, with scores ranging from 1 (totally disagree) to 5 (totally Agree). The goal was to explore if any of these traits could affect the memory recognition for the stimuli. The demographic questionnaire also incorporated the consent form.

A pre-test was developed, in which users would see 20 pictures of real flags (from both countries and U.S. states), with a 3-seconds interval, to be remembered later. The test was developed and administrated using PsychoPy v3.0. The experiment was also developed and administrated with PsychoPy v3.0, showing 60 pictures of faces, scenes, and devices, as explained above.

Design

The test was designed as a within-subject experiment, every participant took part in the same memory recognition task and was exposed to all the stimuli. The order of presentation of the stimuli was randomized.

Procedure

Each participant was tested individually in a silent and comfortable space using a 16’ monitor. Five participants did the test remotely through a screen-sharing application (Skype), due to the limitations imposed by the COVID-19 crisis.

As a first step, a consent form was issued to each participant. Once understood and signed, each participant had to fill out a survey. Directly after this, the participants were assigned to a practice round in which they had to memorise 20 images of flags that were shown with an interval of three seconds. After all the images were shown, the participants had to see 20 images of flags again, of which 50% appeared before and 50 % were new. This time, when each image was shown, the participants had to press ‘Y’ on the keyboard if they thought that they saw it before, and ‘N’ in case they did not. The purpose of this pre-experiment was to measure the memory recognition score in a task with neutral stimuli (Flags), where trustworthiness was not a condition yet, to compare each participant’s score to the scores obtained for untrustworthy and trustworthy stimuli (Basal performance).

Immediately after the pre-test, the participants started with the experiment. They were instructed to memorize a total of 60 images (including faces, scenes, and devices) which were again shown with an interval of three seconds. In this round, 50% of the images portrayed untrustworthy stimuli and 50% portrayed trustworthy stimuli, although the participants were unaware of this categorization. Once all the images were shown, the participants took a 30-minutes break, during which they were advised not to do any task that would use a lot of working memory.

Back from the break, participants were shown again sixty pictures, 50% of which appeared before, and 50 % were new. Like in the pre-experimental phase, participants had to press “Y” if they thought that they saw the picture before, and “N” in case they did not. For the participants that were tested remotely, they had to indicate aloud whether they saw an image, after which the researcher pressed the corresponding key.

When the whole experiment was finished, participants were thanked for their participation in the experiment and offered the possibility to learn about the results of this study when available.

Data Analysis

The data collected through the PsychoPy program was exported in excel and later processed, obtaining a workable data set for analysis in SPSS Statistics. To test the memory recognition performance of each participant, a measure of discrimination (D-Prime score or d') was calculated, determining the participant's ability to discriminate between old and new images. This metric, derived from the techniques of the Signal Detection Theory, provides an explicit metric that expresses the difference between normalized hit and false alarm rates [d'=z(H) -z(F)].

The advantage of this technique is that it provides a unitless measure that considers both response bias (General tendency to respond yes/no), and sensitivity (Degree of overlap between signal and noise distributions) (Stanislaw & Todorov, 1999). Therefore, D-Prime score is unaffected by response bias (Anderson, 2015). Since old and new images are sampled repeatedly, D-Prime score allows to effectively measure the participant's ability to discriminate the signal, which are images presented before, from the noise, i.e., the new images. (Macmillan & Creelman, 2004).

In the case of the pre-test, a single D-Prime score per participant quantified how good each participant was at recognizing previously shown images of flags. For the second round, several D-Prime scores were calculated per each participant, to see if there were differences in memory recognition amongst the two conditions (Trustworthy and Untrustworthy stimuli), for each type of stimuli (Scenes, Devices and Faces). The overall memory recognition performance per condition was also calculated, resulting in a total of 9 D-Prime scores per participant for this round.

The D-Prime score of the participants ranged from below 0 (bad detectors) to above two (particularly good detectors). To study the influence of different personality factors, a Delta score or ∆ (∆= D-Prime score Untrustworthy condition - D-Prime score Trustworthy condition) was calculated to divide the participants who were better at remembering trustworthy stimuli from those who remembered better untrustworthy stimuli.

- BTD (Better Trustworthy Detectors, ∆<0): Participants with a higher D-Prime score for trustworthy than untrustworthy stimuli.
- BUD (Better Untrustworthy Detectors, ∆>0): Participants with higher D-Prime score for untrustworthy than trustworthy stimuli.

This division of participants allowed to create 2 groups: Group 0 (BUD) and group 1 (BTD). By selecting cases, the demographic data for each BTD and BUD was explored independently as to compare the performance and demographic variables of each group independently.

The demographic variables were also used to explore differences in memory recognition. Some of the demographic data (age, gender, and nationality) was obtained directly from the participant's questionnaire, whilst other variables had to be quantified as follows:

- Education was ranked from 1 (less than high school degree) to 8 (professional degree). The groups were divided in 1 (below level 4 or HBO) and 2 (above or at level 4 or HBO).
- Trust score was obtained using a cumulative score (-2 to +2 per each question, with a scale of -12 to +12) and following the guidelines by Yamagishi and Yamagishi (1994).

- Geekism score was obtained using a cumulative score (-2 to +2 per each question, with a scale of -32 to +32) and following the guidelines by Sander (2013).
- Technology acceptances score was obtained using a cumulative score (-2 to +2 per each question, with a scale of -24 to +24) and following the guidelines by Rosen et al. (2013).

Results

Descriptive Statistics

The participants of this experiment were well-educated, with an average level of study corresponding to HBO level (μ=4.31; σ=1.493). On average, the participants had slightly positive results on trust, geekism and technology acceptance personality scales. A table inserted below displays the average values for the demographic variables as well as memory performance per type of stimuli.

Table 1

Average scores of personality scales and memory performance for Pre-Test and for Experimental Test, divided per type of stimuli and subcategory.

Measure	*M*	*SD*	Range	Confidence Interval
Trust Score	2.37	2.82	-6 to 8	1.31 to 3.42
Geekism Score	1.87	8.73	-12 to 17	-1.39 to 5.12
Technology Acceptance Score	3.20	4.99	-9 to 13	1.34 to 5.06
D-Prime score (Pre-Test)	1.35	.47	-.19 to 1.96	1.18 to 1.52
D-Prime score (Trustworthy Stimuli)	1.13	.48	.28 to 2.18	.95 to 1.31
D-Prime score (Untrustworthy Stimuli)	1.14	.50	.20 to 1.95	.95 to 1.32
D-Prime score (Trustworthy Faces)	.69	.56	-.25 to 1.73	.49 to .90
D-Prime score (Untrustworthy Faces)	.60	.57	-.89 to 1.58	.39 to .81
D-Prime score (Trustworthy Scenes)	1.36	.36	.37 to 1.80	1.22 to 1.49
D-Prime score (Untrustworthy Scenes)	.95	.63	-.83 to 1.80	.72 to 1.19
D-Prime score (Trustworthy Devices)	1	.47	.29 to 2.06	.82 to 1.17
D-Prime score (Untrustworthy Devices)	.99	.56	0 to 2.06	.78 to 1.20

The average hit rate in the experiment was 0.41, and the false alarm rate was 0.102 for the trustworthy stimuli, and 0.109 for the untrustworthy stimuli. The measure of discrimination (d') was 1.13 for the trustworthy stimuli and 1.14 for the untrustworthy stimuli. A One-Sample T-test was performed against a test value of 0 (which indicates an inability to distinguish signals from noise) for the d' Trustworthy ($t(29) = 13.0$, $p < .001$) and for d' Trustworthy, ($t(29) = 12.45$, $p < .001$), endorsing that, in general, people were able to correctly recall old images and perform the memory task.

The differences in Discriminability Index (d') amongst the conditions trustworthy and untrustworthy were very small, except for the category "Scenes". This suggests that memory recognition did not differ greatly amongst the conditions of trustworthiness for the categories of Devices and Faces, nor for the total results.

Main Analyses

To test the differences in memory recognition between the trustworthy and untrustworthy conditions, we performed an ANCOVA [within-subjects factor: Trustworthiness (Trustworthy, Untrustworthy); covariate: D-Prime Pre-Test] with repeated measures. In this analysis, we used D-Prime score from the pre-experimental phase as a covariate for all participants and kept memory recognition performance (D-Prime score) as a dependent variable. Results are presented below, in Table 2:

Table2

Analysis of Covariance for overall Memory Recognition performance by trustworthiness condition with Pre-Test Memory Recognition performance as a covariate.

Source	SS	df	MS	F	p	η^2
Performance Pre-Test (Covariate)	.194	1	.194	1.503	.230	.051
Trustworthiness	.208	1	.208	1.615	.214	.055
Error	3.609	28	.129			

Results on the ANCOVA for memory performance on trustworthy images versus untrustworthy images, controlled for pre-test memory performance, indicated no statistically significant differences in memory recognition (*F* 1,28) = 1.615*, p* = .214. [1]Effects on subcategory on memory performance were also not found when studying each category (Scenes, Devices and Faces) separately to see differences in memory recognition between trustworthy and untrustworthy stimuli.

As mentioned in the methods section, participants were divided in: Better Untrustworthy Detectors (BUD) and Better Trustworthy Detectors (BTD). This was aimed to explore differences between participants who were better at remembering trustworthy stimuli and participants who were better at remembering untrustworthy stimuli on some personality and demographic factors. Means and standard deviation for each group are reported below.

Table 3

Comparison on demographic values of Better Trustworthy Detectors (BTD) and Better Untrustworthy Detectors (BUD) for the general set of stimuli (Faces, Scenes & Devices).

	N BTD	BTD Mean	S.D.	Confidence Interval	N BUD	BUD Mean	S.D.	Confidence Interval
Age	16	27.13	7.82	22.96 to 31.29	14	33.14	13.86	25.14 to 41.15
Trust Score	16	2.94	2.98	1.35 to 4.52	14	1.71	2.59	.22 to 3.21
Geekism Score	16	2.69	9.44	-2.34 to 7.72	14	. 93	8.07	-3.73 to 5.59
Technology Acceptance	16	3.19	5.91	.04 to 6.34	14	3.21	3.89	.97 to 5.46

A t-Test was performed for each of these variables in order to test the significance of the differences in demographic values between BUD and BTD.

- Age was not significantly different for BUD and BTD, *t* (28) =-1.489, *p*=.148.

[1] A paired-samples t-test was conducted to compare the memory performance for untrustworthy and trustworthy stimuli. There was no significant difference in the memory recognition for trustworthy stimuli (M=1.13, SD=.48) and untrustworthy stimuli (M=1.14, SD=.50); t(29)=-.07, p=.945. A significant difference was observed in the memory recognition for trustworthy scenes (M=1.36, SD=.36) and untrustworthy scenes (M=.95, SD=.63); t (29) =4.072, p <.001). No significant differences were found in the case of faces and devices.

- Trust score was not significantly different for BUD and BTD, *t* (28) = 1.193, *p*= .243.
- Geekism score was not significantly different for BUD and BTD, *t* (28) = .544, *p*= .591.
- Technology Acceptance score was not significantly different for BUD and BTD, *t* (28) = -.014, *p*= .989.

Table 4

Comparison on demographic values of Better Trustworthy Detectors (BTD) and Better Untrustworthy Detectors (BUD) for the category of Devices.

	N BTD	BTD Mean	S.D.	Confidence Interval	N BUD	BUD Mean	S.D.	Confidence Interval
Age	16	30.5	11.58	24.33 to 36.67	14	29.29	11.32	25.14 to 41.15
Trust Score	16	2.94	3.19	1.24 to 4.64	14	1.71	2.27	.22 to 3.21
Geekism Score	16	3.31	9.27	-1.63 to 8.25	14	. 21	8.07	-3.73 to 5.59
Technology Acceptance	16	4.44	4.44	2.07 to 6.80	14	1.79	5.35	.97 to 5.46

A t-Test was performed for each of these variables in order to test the significance of the differences in demographic values between BUD and BTD of pictures of devices.

- Age was not significantly different for BUD and BTD, *t* (28) =.29, *p*=.774.
- Trust score was not significantly different for BUD and BTD, *t* (28) = 1.193, *p*= .243.
- Geekism score was not significantly different for BUD and BTD, *t* (28) = .969, *p*= .341.
- Technology Acceptance score was not significantly different for BUD and BTD, *t* (28) = 1.483, *p*= .149. Further investigation is needed to explore this trend effect that points towards a possible relationship between technology acceptance and trust towards systems.

Discussion

A memory test was carried out, expecting people to remember better a set of untrustworthy stimuli, compared to a set of trustworthy stimuli. Although the participants of this study showed an ability to discriminate new from old pictures, this ability was, in general, independent of the subcategory of the stimuli (Trustworthy or Untrustworthy). In the case of scenes, the images of trustworthy scenes were remembered better than the images of untrustworthy scenes. These results contradicted the assumptions based on previous studies of enhanced ability of people to remember untrustworthy stimuli (Yamagishi et al., 2003; Chiappe et al., 2004; Verplaetse et al., 2007; Bell & Buchner, 2007).

By looking at demographic and personality characteristics (i.e., Trust, Geekism and technology acceptance), the results suggested that participants who performed better at remembering untrustworthy stimuli did not differ significantly from participants who were better detectors of trustworthy stimuli. This is quite in line with the findings of Oda (1997), who did not find differences in recognition of co-operators and cheaters per gender of the participant, despite previous research proposed that females would be better than males at face recognition tasks (McKelvie; 1981; Nesse et al., 1990; Rodin 1987).

With regards to the personality factors, none of them differed significantly between Better Trustworthy Detectors and Better Untrustworthy Detectors, of stimuli in general, as well as between BTD and BUD of images of devices A possible trend suggesting that better detectors of trustworthy devices might have a higher Technology Acceptance score compared to better detectors of untrustworthy devices was identified, with no significant relevance. Further investigation is needed to study this trend effect, extending the scope of the studies from Gefen et al. (2003) about the effect of technology acceptance on trust towards systems.

The results of this experiment align with previous studies which cast doubt upon the idea of an enhanced memory recognition for faces of cheaters. Suzuki et al. (2010) argued that facial trustworthiness has limited predictive power, due to some biases that influence the assessment of trustworthiness, such as a resemblance to one's own face and emotional expression. In addition to this, Buchner et al. (2009) claimed that recognizing a face as already seen is not useful per se in avoiding cheaters and has, therefore, no adaptive significance.

Furthermore, as observed by various authors (Brown & Moore, 2002; Frank et al., 1993; Verplaetse et al., 2007), it is unlikely that predictive detection of cheaters can rely on permanent features showed on still photographs. Instead of permanent cues, Yamagishi et al. (2003) proposed that emotional expressions or gestures could be necessary for the advanced memory for faces of cheaters, as it was the case in the experiment of Verplaetse et al. (2007).

Lastly, the attractiveness of the stimuli was not measured, contrary to the experiments of Mealey (1996) and Oda (1997), who proposed that a higher attractiveness could be related to a higher memory recognition. This might be linked to the increased memory for the images of trustworthy scenes, as these were more aesthetically pleasing than the images for untrustworthy scenes.

Implications

The results of the memory test showed that people did not have an enhanced memory recognition for pictures of untrustworthy devices. Rather opposite, in the present experiment participants exhibited a higher memory recognition performance for the trustworthy condition, in each of the 3 categories of stimuli, although these differences were not statistically significant overall, (only for the category "scenes"). These results were not aligned with the experiment by Verplaetse et al. (2007), in which it was found that participants had an enhanced memory towards the images of untrustworthy faces.

Generally speaking, the results of this exploratory experiment would not fit with the theory of an adapted cognitive mechanism that evolves to deal with untrustworthy devices. Conversely, the results are in line with Barclay and Lalumière (2006) and Mehl and Buchner (2008), who found that memory recognition for faces associated with a history of cheating was not better than memory recognition for faces associated with a history of trustworthiness. A later study by Barclay (2008) suggested that memory differences could be caused by the salience or rarity of the stimuli, rather than by the trustworthiness, leaving room for further research using this variable.

Replications are needed to further investigate the trend effect identified in this research, which could suggest that participants who are better detectors of trustworthy devices have a higher technology acceptance score compared to participants who are better able to detect untrustworthy devices.

Limitations

The generalization of the results from this exploratory research are limited by some factors:
To start with, and as already mentioned, all the pictures used in this experiment were still photographs which showed no emotion or context associated. As reported by Verplaetse et al. (2007), the enhanced memory recognition for faces of cheaters only happened with pictures taken in the proper round of a one-shot prisoner dilemma game, not with neutral-expression and practice-round pictures.

Complementary to this, and unlike most of the previous studies on enhanced memory for cheaters, there was no context of social interaction in this experiment. If a cheater detection is activated only in a situation of social exchange (i.e., cooperation), the methodology of this experiment might have failed to properly activate the underlying cognitive mechanism responsible for enhanced memory towards untrustworthy stimuli. Consequently, and as Van Lier et al. (2013) stated, when the module is not activated, memory performance facilitation should not be expected.

Hence, these results have low ecological validity, as there was a total absence of perceived risk and attached to the memory performance, and the experiment setting did not aim to simulate any setting of the day-to-day life. In addition to this, the fact that there was no compensation (Either monetary or academic) for the participants could arguably have affected the performance of the participants. As Yamagishi et al. (2003) mentioned, making the total amount dependant on how many pictures are correctly remembered motivates participants to remember as many items as possible. In general, it could be argued that participants did not have sufficient extrinsic motivation to perform at the best of their capacities, neither a reward, nor the benefit of avoiding a potential risk.

Another possible limitation of this study concerns the time of the break used in the memory experiment, which was of 30 minutes. Although this break is in line with Chiappe et al. (2004), who found that the bias in remembering cheaters was evident even with a short break, the effect of time on memory recognition could not be crosschecked with other time intervals. Also, some studies on memory recognition argue that memory consolidates after a night of sleep, originating from the reactivation of newly encoded memory representations (Rasch & Born, 2013). Gais (2006) found that memory recognition after a night of sleep was enhanced after a night of sleep. It was therefore not possible to explore if the differences in recognition between trustworthy and untrustworthy stimuli got larger over time or remained non-significant.

Additionally, given the exceptional circumstances derived from the COVID-19 and the impossibility to meet some people face-to-face, some of the participants (a total of 5) were tested remotely. Although the results from these participants did not differ from the average total results, meaning it was not necessarily a limitation, it would definitively be best to ensure that the same procedure and materials are used with all the participants to avoid possible effects on the performance.

Recommendations

Several adjustments can be applied to deal with the limitations stated. First, and regarding the activation of the proper cognitive mechanisms associated with memory for cheaters, future research could use images or videos shot of a real context (e.g., a device being operated by a person to perform a task) so that participants will witness how devices look in a real context. This could be combined with new tasks (e.g., trying to find the best value for money from a portfolio of products) followed by a memory recognition exercise in which participants are unexpectedly asked which images correspond to the devices that they have already seen during the primary task.

Applying these changes could discard the limitation of a low ecological validity caused by the usage of still photographs without a context or task of cooperation, which might block the enhanced memory for untrustworthy stimuli. In addition to this, it could be advisable for future research to allocate some form of compensation to the participants to strengthen their motivation.

Using the PsychoPy experiment developed for this work, future researchers could also crosscheck whether the memory recognition was different applying different pause times. Researchers could therefore implement longer pause periods to better test the memory recognition performance, like in the experiments of Verplaetse et al. (2007) or Mealet et al. (1996), whose pause time was 1 hour and 1 week, respectively. In addition to this, researchers might also check whether the results are consistent when more information regarding the product is given. Adding information regarding the device's features or reliability would give more hints to the users to assess the trustworthiness of devices, an approach closer to the methodology applied by Mealey et al. (1996) and Oda (1997), so this could give some insights into people's assessment of devices trustworthiness.

Apart from working out with the limitations, there are some further recommendations that future research could apply. As it has been stated, trust towards systems can also be measured towards other measures, both direct and indirect. Therefore, a questionnaire could be added, in which participants would rate how trustworthy each device seems to them (Direct measurement), and how willing they would be to purchase it (Indirect, Behavioural measure). In this way, different measures of trust towards stimuli could be analysed together under the same study, providing interesting data to evaluate which of these measures is more appropriate in the context of Devices.

Furthermore, this questionnaire could also be used to ask participants to rank the attractiveness of the showed stimuli, which would allow this variable to be controlled. This would allow studying the effect of attractiveness in memory recognition, since high attractiveness might be correlated to better recognition, as Oda (1997) proposed. Alternatively, it could also register people's impressions of how particular the design of different devices is, as the rarity or salience of stimuli could be associated with a higher memory recognition for it (Barclay, 2006).

A particularly interesting hint for future research is the effect of Technology Acceptance on recognition of untrustworthy devices. As the difference between BUD and BTD did not reach significance levels, perhaps due to an insufficient number of participants, it would be recommended for future research to include this question in their experiments, in order to underpin the possible relations amongst these variables. This could also be strengthened by using stratified sampling to obtain a wider range of participants that differ in the factors mentioned above.

Although the present research did not provide evidence that people would be able to identify untrustworthy devices based on visual cues, it is possible that future research will succeed to do so after properly adjusting for the limitations stated above. In that case, the next step would be to pinpoint which features of a design are more looked at during the presentation of the pictures. Further studies can address this by adding complementary tools (such as eye trackers) to learn about the possible differences in stimuli inspection/visualization. Such a tool has already been used by Pan et al. (2007) in their study about trust in the usage of search engines.

Conclusion

This research aimed to test whether there was a difference in remembering images of trustworthy or untrustworthy stimuli, to explore the role of memory in the recognition of what is trustworthy. In line with the theories of enhanced memory for faces of cheaters (Verplaetse et al., 2007), it was expected that participants would remember images of untrustworthy stimuli better than images of trustworthy stimuli. Nonetheless, the participants of the present experiment did not show a difference in remembering untrustworthy or trustworthy stimuli. As some of the participants remembered untrustworthy stimuli better than trustworthy stimuli, it was studied whether any personality or demographic factors could explain this difference, but nothing was found.

This study was limited by the low ecological validity of the task and stimuli, the lack of compensation to the participants, and the usage of a single break condition. By resolving these limitations (e.g., offering compensation to participants, having longer breaks or employing videos of devices showing how they are operated in a real setting), future research can increase the relevance and validity of their outcomes. This could also allow to further investigate the (non-significant) trend identified in this study, which could suggest that participants who are better at remembering faces of trustworthy devices might have a higher technology acceptance compared to participants who are better at remembering untrustworthy devices.

Given the increasing presence and influence of technology in human life, further research addressing people’s ability to identify untrustworthy technology should consider alternative ways of studying cognition and trust towards technology. For instance, other aspects of cognition (e.g., decision-making, attention, and judgement) could be investigated in combination or substitution with memory. This could be accomplished by combining direct measures (Such as rating the trustworthiness of a device) as well as indirect measures (Such as measuring predisposition to use/purchase a particular device. This would provide different references to study Trust Towards Systems in general, and recognition of untrustworthy devices in particular.

References

Andersen, M. (2016, November 2). Deceptive Design is Illegal now, so why are you still getting swindled? *Eye on Design*. https://eyeondesign.aiga.org/deceptive-design-is-illegal-now-so-why-are-you-still-getting-swindled/.

Anderson, N. D. (2015). Teaching signal detection theory with pseudoscience. *Frontiers in Psychology*, 6. https://doi.org/10.3389/fpsyg.2015.00762

Antonak, R. F., & Livneh, H. (1995). Direct and indirect methods to measure attitudes toward persons with disabilities, with an exegesis of the error-choice test method. *Rehabilitation Psychology*, 40(1), 3–24. https://doi.org/10.1037/0090-5550.40.1.3

Ba, S., & Zhang, H. (2003). Building trust in Online auction Markets through an economic incentive mechanism. *Decision Support Systems*, 35(3), 273-286. doi: 10.1016/S0167-9236(02)00074-X

Barclay, P. (2008). Enhanced recognition of defectors depends on their rarity. *Cognition*, 107(3), 817–828. https://doi.org/10.1016/j.cognition.2007.11.013

Barclay, P., & Lalumière, M. L. (2006). Do people differentially remember cheaters? *Human Nature*, 17, 98–113. https://doi.org/10.1007/s12110-006-1022-

Beldad, A., de Jong, M., & Steehouder, M. (2010). How shall I trust the faceless and the intangible? A literature review on the antecedents of online trust. *Computers in Human Behavior*, 26(5), 857–869. doi:10.1016/j.chb.2010.03.013

Bell, R., & Buchner, A. (2009). Enhanced Source Memory for Names of Cheaters. *Evolutionary Psychology*, 7(2). https://doi.org/10.1177/147470490900700213´

Bell, R., & Buchner, A. (2012). How Adaptative Is Memory for Cheaters? Current directions in Psychological Science, 21(6), 403-408. https://doi.org/ 10.1177/0963721412458525.

Benbasat, I., & Wang, W. (2005). Trust in and Adoption of Online Recommendation Agents*. Journal of the Association for Information Systems*, 6(3), 72–101. https://doi.org/10.17705/1jais.00065

Borsci, S., Kuljis, J., Barnett, J., & Pecchia, L. (2014). Beyond the User Preferences: Aligning the Prototype Design to the Users' Expectations. *Human Factors and Ergonomics in Manufacturing & Service Industries*, 26(1), 16–39. https://doi.org/10.1002/hfm.20611

Borsci, S., Uchegbu, I., Buckle, P., Ni, Z., Walne, S., & Hanna, G. B. (2017). Designing medical technology for resilience: integrating health economics and human factors approaches. Expert Review of Medical Devices, 15(1), 15–26. https://doi.org/10.1080/17434440.2018.1418661

Borsci, S., Buckle, P., Walne, S., & Salanitri, D. (2018). Trust and Human Factors in the Design of Healthcare Technology. In S. Bagnara, R. Tartaglia, S. Albolino, T. Alexander, & Y.Fujita (Eds.), *Proceedings of the 20th Congress of the International Ergonomics Association (IEA 2018): Volume VII: Ergonomics in Design, Design for All, Activity Theories for Work Analysis and Design, Affective Design (Vol. 824, pp. 207-215*). (Advances in Intelligent Systems and Computing; Vol. 824). Springer

Brignull, H. (2013, August 29). Dark Patterns: inside the interfaces designed to trick you. *The Verge*. https://www.theverge.com/2013/8/29/4640308/dark-patterns-inside-the-interfaces-designed-to-trick-you

Brysbaert, M. (2019). How many participants do we have to include in properly powered experiments? A tutorial of power analysis with reference tables. *Journal of Cognition*, 2(1), 16. http://doi.org/10.5334/joc.72

Brown, W., Moore, C. (2002). Smile Asymmetries and reputation as reliable indicators of likelihood to cooperate: An evolutionary analysis. *Advances in psychology research*, 11, 59-78.

Brunswick, G. J. (2014). A Chronology of The Definition of Marketing. *Journal of Business & Economics Research (JBER)*, 12(2), 105 - 114. https://doi.org/10.19030/jber.v12i2.8523.

Buchner, A., Bell, R., Mehl, B., & Musch, J. (2009). No enhanced recognition memory, but better source memory for faces of cheaters. *Evolution and Human Behavior*, 30(3), 212–224. https://doi.org/10.1016/j.evolhumbehav.2009.01.004

Chiappe, D., Brown, A., Dow, B., Koontz, J., Rodriguez, M., & McCulloch, K. (2004). Cheaters Are Looked at Longer and Remembered Better than Cooperators in Social Exchange Situations. *Evolutionary Psychology*, 2, 108 - 120. https://doi.org/10.1177/147470490400200117

Cook, G. I., Marsh, R. L., & Hicks, J. L. (2003). Halo and devil effects demonstrate valenced-based influences on source-monitoring decisions. *Consciousness and Cognition: An International Journal*, 12(2), 257–278. https://doi.org/10.1016/s1053-8100(02)00073-9

Cosmides, L. (1989). The logic of social exchange: Has natural selection shaped how humans reason? Studies with the Watson selection task. *Cognition*, 31 (3),187–276. https://doi.org/10.1016/0010-0277(89)90023-1

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.

Cosmides, L., & Tooby, J. (2005). Neurocognitive adaptations designed for social exchange. In D. M. Buss (Ed.), *The Handbook of evolutionary psychology* (pp. 584-627). John Wiley & Sons, Inc.

Cosmides, L., Barrett, H. C., & Tooby, J. (2010). Adaptive specializations, social exchange, and the evolution of human intelligence. *Proceedings of the National Academy of Sciences of the United States of America*, 107(2), 9007–9014. Doi: 10.1073/pnas.0914623107

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), Article e0190954. https://doi.org/10.1371/journal.pone.0190954

Das, T. K., & Teng, B. S. (2004). The risk-based view of trust: A conceptual framework. *Journal of Business and Psychology*, 19(1), 85–116. Doi: 10.1023/B:JOBU.0000040274.23551.1B

De Angeli, A., Sutcliffe, A., & Hartmann, J. (2006). Interaction, usability and aesthetics: What influences users' preferences? *Proceedings of the 6th Conference on Designing Interactive Systems* (pp. 271–280). University Park, PA: ACM.

Ermer, E., Guerin, S. A., Cosmides, L., Tooby, J., & Miller, M. B. (2006). Theory of mind broad and narrow: Reasoning about social exchange engages ToM areas, precautionary reasoning does not. *Social Neuroscience*, 1(3–4), 196–219. https://doi.org/10.1080/17470910600989771

Ermer, E., Cosmides, L., Tooby, J. (2007). Cheater Detector Mechanism. In: *Encyclopedia of Social Psychology*. Thousand Oaks, CA: SAGE Publications

Farrelly, D., & Turnbull, N. (2008). The Role of Reasoning Domain on Face Recognition: Detecting Violations of Social Contract and Hazard Management Rules. *Evolutionary Psychology*, 6(3), 523 - 537. https://doi.org/10.1177/147470490800600317

Frank, R. H., Gilovich, T., & Regan, D. T. (1993). The evolution of one-shot cooperation: An experiment. *Ethology and Sociobiology*, 14(4), 247-256. https://doi.org/10.1016/0162-3095(93)90020-I

Fogg, B. J., Soohoo, C., Danielson, D. R., Marable, L., Stanford, J., & Tauber, E. R. (2003). How do users evaluate the credibility of Web sites? *Proceedings of the 2003 Conference on Designing for User Experiences - DUX '03*. doi:10.1145/997078.997097

Gais, S. (2006). Sleep after learning aids memory recall. *Learning & Memory*, 13(3), 259–262. doi:10.1101/lm.132106.

Gefen, D. (2000). E-commerce: the role of familiarity and trust. *Omega*, 28(6), 725–737. doi:10.1016/s0305-0483(00)00021-9

Gefen, D., Karahanna, E., & Straub, D. (2003). Trust and TAM in Online Shopping: An Integrated Model. *MIS Quarterly*, 27(1), 51 - 90. doi:10.2307/30036519

Glaeser, E. L., Laibson, D. I., Scheinkman, J. A., & Soutter, C. L. (2000). Measuring Trust. *The Quarterly Journal of Economics*, 115(3), 811–846. https://doi.org/10.1162/003355300554926

Gol Mohammadi N. (2019) Trustworthiness-by-design. In: *Trustworthy Cyber-Physical Systems*. Springer Vieweg, Wiesbaden. https://doi.org/10.1007/978-3-658-27488-7_6

Greenberg, S., Boring, S., Vermeulen, J., Dostal, J. (2014). Dark patterns in proxemic interactions: a critical perspective. *Proceedings of the 2014 conference on designing interactive systems - DIS'14*. doi:10.1145/2598510.2598541

Hale, J., Payne, M. E., Taylor, K. M., Paoletti, D., & De C Hamilton, A. F. (2018). The virtual maze: A behavioural tool for measuring trust. *Quarterly Journal of Experimental Psychology*, 71(4), 989–1008. https://doi.org/10.1080/17470218.2017.1307865

Harley, A. (2016, May 8). Trustworthiness in Web Design: 4 Credibility Factors. *NN Group*. https://www.nngroup.com/articles/trustworthy-design/

Herse, S., Vitale, J., Tonkin, M., Ebrahimian, D., Ojha, S., Johnston, B., Williams, M.A. (2018). Do You Trust Me, Blindly? Factors Influencing Trust Towards a Robot Recommender System. *2018 27th IEEE International Symposium on Robot and Human Interactive Communication* https://doi:10.1109/roman.2018.8525581

Hosmer, L. T. (1995). TRUST: THE CONNECTING LINK BETWEEN ORGANIZATIONAL THEORY AND PHILOSOPHICAL ETHICS. *Academy of Management Review*, 20(2), 379–403. https://doi.org/10.5465/amr.1995.9507312923

Huvila, I. (2017). Distrust, mistrust, untrust and information practices. Information Research, 22(1)

Khalighy, S., Green, G., Scheepers, C., & Whittet, C. (2015). Quantifying the qualities of aesthetics in product design using eye-tracking technology. *International Journal of Industrial Ergonomics*, 49, 31–43. doi:10.1016/j.ergon.2015.05.011

Kovács-Bálint, Z., Bereczkei, T., & Hernádi, I. (2013). The telltale face: possible mechanisms behind defector and cooperator recognition revealed by emotional facial expression metrics. *British journal of psychology*), 104(4), 563–576. https://doi.org/10.1111/bjop.12007

Lankton, N., McKnight, D. H., & Tripp, J. (2015). Technology, Humanness, and Trust: Rethinking Trust in Technology. *Journal of the Association for Information Systems*, 16(10), 880–918. https://doi.org/10.17705/1jais.00411

Lee, M.K.O., Turban, E. (2001). A trust model for consumer internet shopping. *International Journal of Electronic Commerce* 6 (1), 75–91. https://doi.org/10.1080/10864415.2001.11044227

Li, X., Hess, T. J., & Valacich, J. S. (2008). Why do we trust new technology? A study of initial trust formation with organizational information systems. *Journal of Strategic Information Systems*, 17(1), 39–71. https://doi.org/10.1016/j.jsis.2008.01.001

Lippert, S. K., & Michael Swiercz, P. (2005). Human resource information systems (HRIS) and technology trust. *Journal of Information Science*, 31(5), 340–353. https://doi.org/10.1177/0165551505055399

Ma, D. S., Correll, J., & Wittenbrink, B. (2015). The Chicago face database: A free stimulus set of faces and norming data. *Behavior Research Methods*, 47(4), 1122–1135. https://doi.org/10.3758/s13428-014-0532-5

Macmillan, N. A. (2002). Signal detection theory. In H. Pashler & J. Wixted (Eds.), Stevens' handbook of experimental psychology: Methodology in experimental psychology (p. 43–90). John Wiley & Sons Inc.

Macmillan, N. A., & Creelman, D. C. (2004*). Detection Theory: A User's Guide* (2nd ed.). Psychology Press.

Mayer, R. C., Davis, J. H., & Schoorman, F. D. (1995). An Integrative Model of Organizational Trust. *Academy of Management Review*, 20(3), 709–734. https://doi.org/10.5465/amr.1995.9508080335

Marsh, S., Dibben, M.R. (2005). Trust, Untrust, Distrust and Mistrust – An exploration of the Dark(er) side. *Lecture Notes in Computer Science*. 3477,17-33. doi: 10.1007/11429760_2

McKelvie, S. J. (1981). Sex differences in memory for faces. *The Journal of Psychology: Interdisciplinary and Applied*, 107(1), 109–125. https://doi.org/10.1080/00223980.1981.9915211

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*, 2(2), 1–25. https://doi.org/10.1145/1985347.1985353

McKnight, H.D. (2005). Trust in Information Technology. In G. B. Davis (Ed.) *The Blackwell Encyclopaedia of Management Vol. 7 Management Information Systems* (pp.329-331). Minnesota: Blackwell

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. https://doi.org/10.1287/isre.13.3.334.81

McKnight, D. H., Cummings, L. L., & Chervany, N. L. (1998). Initial Trust Formation in New Organizational Relationships. *The Academy of Management Review*, 23(3), 473 - 490. https://doi.org/10.2307/259290

Mealey, L., Daood, C., & Krage, M. (1996). Enhanced memory for faces of cheaters. *Ethology and Sociobiology*, 17(2), 119–128. https://doi.org/10.1016/0162-3095(95)00131-x

Mehl, B., & Buchner, A. (2008). No enhanced memory for faces of cheaters. *Evolution and Human Behavior*, 29(1), 35–41. https://doi.org/10.1016/j.evolhumbehav.2007.08.001

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

Nesse, R.M., Silverman, A., Bortz, A. (1990). Sex Differences in Ability to Recognize Family Resemblance. Ethology & Sociobiology, 11 (1), 11 - 21. https://doi.org/10.1016/0162-3095(90)90003-O

Nickel, P. J., Franssen, M., & Kroes, P. (2010). Can We Make Sense of the Notion of Trustworthy Technology? *Knowledge, Technology & Policy*, 23(3–4), 429–444.

Oda, R. (1997). Biased face recognition in the Prisoner's Dilemma Game. *Evolution and Human Behavior*, 18(5), 309–315. https://doi.org/10.1016/S1090-5138(97)00014-7

Oosterhof, N. N., & Todorov, A. (2009). Shared perceptual basis of emotional expressions and trustworthiness impressions from faces. *Emotion*, 9(1), 128–133. https://doi.org/10.1037/a0014520

Pan, B., Hembrooke, H., Joachims, T., Lorigo, L., Gay, G., & Granka, L. (2007). In Google We Trust: Users' Decisions on Rank, Position, and Relevance. *Journal of Computer-Mediated Communication*, 12(3), 801–823. https://doi.org/10.1111/j.1083-6101.2007.00351.x

Pengnate, S., & Sarathy, R. (2017). An experimental investigation of the influence of website emotional design features on trust in unfamiliar online vendors. *Computers in Human Behavior*, 67, 49–60. doi:10.1016/j.chb.2016.10.018.

Rasch, B., & Born, J. (2013). About Sleep's Role in Memory. *Physiological Reviews*, 93(2), 681–766. doi:10.1152/physrev.00032.2012

Rodin, M. J. (1987). Who Is Memorable to Whom: A Study of Cognitive disregard. *Social Cognition*, 5(2), 144–165. https://doi.org/10.1521/soco.1987.5.2.144

Salam, A. F., Iyer, L., Palvia, P., & Singh, R. (2005). Trust in e-commerce. *Communications of the ACM*, 48(2), 72–77. doi:10.1145/1042091.1042093

Salanitri, D., Hare, C., Borsci, S., Lawson, G., Sharples, S., & Waterfield, B. (2015). Relationship between trust and usability in virtual environments: An ongoing study. In M. Kurosu (Ed.), Human-Computer Interaction: Design and Evaluation. HCI 2015. Lecture Notes in Computer Science, Vol. 9169. Springer, Cham. https://doi.org/10.1007/978-3-319-20901-2_5

Shneiderman, B. (2000). Designing trust into online experiences. *Communications of the ACM*, 43(12), 57-59. https://dl.acm.org/doi/10.1145/355112.355124

Stanislaw, H., & Todorov, N. (1999). Calculation of signal detection theory measures. *Behavior Research Methods, Instruments, & Computers*, 31(1), 137–149. https://doi.org/10.3758/BF03207704

Sterling, B. (2014, February 10). Dark Patterns: user interfaces designed to trick people. *Wired*. https://www.wired.com/2014/02/dark-patterns-user-interfaces-designed-trick-people/.

Suzuki, A., & Suga, S. (2010). Enhanced memory for the wolf in sheep's clothing: Facial trustworthiness modulates face-trait associative memory. Cognition, 117(2), 224–229. https://doi.org/10.1016/j.cognition.2010.08.004

Tan,Y., Thoen, W. (2014). Toward a Generic Model of Trust for Electronic Commerce. *International Journal of Electronic Commerce*, 5(2), 61-74. https://doi.org/10.1080/10864415.2000.11044201

Teo, T.S.H. & Liu, J. (2007). Consumer trust in e-commerce in the United States, Singapore and China. *Omega*, 35(1), 22–38. doi:10.1016/j.omega.2005.02.001

Thatcher, J. B., McKnight, D. H., Baker, E. W., Arsal, R. E., & Roberts, N. H. (2011). The Role of Trust in Postadoption IT Exploration: An Empirical Examination of Knowledge Management Systems. *IEEE Transactions on Engineering Management*, 58(1), 56–70. https://doi.org/10.1109/tem.2009.2028320

Todorov, A., Pakrashi, M., & Oosterhof, N. N. (2009). Evaluating Faces on Trustworthiness After Minimal Time Exposure. *Social Cognition*, 27(6), 813–833. https://doi.org/10.1521/soco.2009.27.6.813

Tractinsky, N., Katz, A. S., & Ikar, D. (2000). What is beautiful is usable. *Interacting with Computers*, 13(2), 127–145. https://doi.org/10.1016/S0953-5438(00)00031-X

Van Lier, J., Revlin, R., & De Neys, W. (2013). Detecting Cheaters without Thinking: Testing the Automaticity of the Cheater Detection Module. *PLoS ONE*, 8(1), e53827. doi:10.1371/journal.pone.0053827

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. https://doi.org/10.1016/j.evolhumbehav.2007.04.006

Wixom, B. H., & Todd, P. A. (2005). A Theoretical Integration of User Satisfaction and Technology Acceptance. *Information Systems Research*, 16(1), 85–102. https://doi.org/10.1287/isre.1050.0042

Yamagishi, T., & Yamagishi, M. (1994). Trust and commitment in the United States and Japan. *Motivation and emotion*, 18 (2), 129-166. https://doi.org/10.1007/BF02249397

Yamagishi, T., Tanida, S., Mashima, R., Shimoma, E., & Kanazawa, S. (2003). You can judge a book by its cover: Evidence that cheaters may look different from cooperators. *Evolution and Human Behavior*, 24(4), 290–301. https://doi.org/10.1016/S1090-5138(03)00035-7

Appendixes

Appendix A. Questionnaire.

Survey Flow

Block: Demographics & Education (5 Questions)
Standard: General trust scale (1 Question)
Standard: Geekism scale (1 Question)
Standard: Attitudes towards technology/ Dependency on technology (1 Question)

Page Break

Start of Block: Demographics & Education

Q3 What is your age?

____

Q5 What is your sex?

○ Male (1)

○ Female (2)

○ Other (3)

Q7 What is your nationality?

____

Q6 What is the highest level of school you have completed or the highest degree you have received?

○ Less than high school degree (1)

○ High school graduate (high school diploma or equivalent including GED) (2)

○ Some college but no degree (3)

○ Associate degree in college (2-year) (4)

○ Bachelor's degree in college (4-year) (5)

○ Master's degree (6)

○ Doctoral degree (7)

○ Professional degree (JD, MD) (8)

Q6 What is your occupation/ What do you study?

____

End of Block: Demographics & Education

Start of Block: General trust scale

Q9 Using the following scale, please indicate how much you agree or disagree with the following statements.

	Strongly disagree (1)	Disagree (2)	Neither Agree nor Disagree (3)	Agree (4)	Strongly Agree (5)
Most people are basically honest (1)	○	○	○	○	○
Most people are trustworthy (2)	○	○	○	○	○
Most people are basically good and kind (3)	○	○	○	○	○
Most people are trustful of others (4)	○	○	○	○	○
I am trustful (5)	○	○	○	○	○
Most people will respond in kind when they are trusted by others (6)	○	○	○	○	○

End of Block: General trust scale

Start of Block: Geekism scale

Q10 Using the following scale, please indicate how much you agree/disagree with the following statements.

	Strongly Disagree (1)	Disagree (2)	Neither Agree nor Disagree (3)	Agree (4)	Strongly Agree (5)
I want to understand how computers/software work. (1)	○	○	○	○	○
When someone needs help with a computer, I do my best to help them (2)	○	○	○	○	○
When it comes to computers and internet, privacy settings are very important to me (3)	○	○	○	○	○
Complex tasks with technical devices put me off. (4)	○	○	○	○	○
I once used a technical device for other than the intended purposes (5)	○	○	○	○	○
Objectivity is important to me (6)	○	○	○	○	○
I have the feeling that I don't have much control over my technical devices (7)	○	○	○	○	○
In my free time I don't spend more time with technical devices than other people (8)	○	○	○	○	○
When I purchase a technical device performance is more important to me than looks (9)	○	○	○	○	○
It motivates me to modify technical devices to make them fit my needs (10)	○	○	○	○	○

I already published my own project(s) on the internet or would do it (11)	○	○	○	○	○
I think there are people that would call me a computer freak (12)	○	○	○	○	○
The programming of software does not interest me (13)	○	○	○	○	○
I avoid the advanced features of my technical devices (14)	○	○	○	○	○
I like to share my ideas and projects with others (15)	○	○	○	○	○
Challenging tasks with technical devices excite me (16)	○	○	○	○	○

End of Block: Geekism scale

Start of Block: Attitudes towards technology/ Dependency on technology

	Strongly Disagree (1)	Disagree (2)	Neither Agree nor Disagree (3)	Agree (4)	Strongly Agree (5)
I feel it is important to be able to find any information whenever I want online (1)	○	○	○	○	○
I feel it is important to be able to access the Internet any time I want (2)	○	○	○	○	○
I think it is important to keep up with the latest trends in technology (3)	○	○	○	○	○
I get anxious when I don’t have my cell phone (4)	○	○	○	○	○
I get anxious when I don’t have the Internet available to me (5)	○	○	○	○	○
I am dependent on my technology (6)	○	○	○	○	○
Technology will provide solutions to many of our problems (7)	○	○	○	○	○
With technology anything is possible (8)	○	○	○	○	○
I feel that I get more accomplished because of technology (9)	○	○	○	○	○
New technology makes people waste too much time (10)	○	○	○	○	○
New technology makes life more complicated (11)	○	○	○	○	○
New technology makes people more isolated (12)	○	○	○	○	○

End of Block: Attitudes towards technology/ Dependency on technology

Appendix B. List of Devices.

Type	Untrustworthy	Source	Trustworthy
Hearing Aids	Nano CIC hearing aids	hearingaidknow.com	Rexton Inox 40 6C
Nebulizer system	Aquilon Medical Nebulisers	U.K. government	Philips InnoSpire Mini
Smartwatch	Basis Peak Watch	Company recalls	Fitbit Charge 4
Toys	Flying tiger Copenhagen Toy Train Charts	CPSC	Kidzzy Toys Train Set
Stairlifts	Acorn 180 Curved Stairlift	CPSC	Handicare Simplicity Plus
Toaster	Bodum Bistro Toasters	CPSC	Breville BTA730XL
Life Alert System	Lively Mobile Plus Emergency Alert Devices	CPSC	LifeStation Mobile 4G LTE
Space Heaters	Amazon Basics 1500-Watt Ceramic Space Heaters	CPSC	Lasko 5409
Wireless Charging	Spansive Source wireless	CPSC	heyday™ Qi
Alarm	"-IV" Intelligent photoelectric smoke sensors	CPSC	First Alert SCO5CN
Baby Food Maker	Babycook Neo steam cooker/blenders	CPSC	Infanso BF300
Baby Monitor	Gynoii GPW-1025	Reviews	Hello Baby HB32
Baby Crib	DaVinci Cribs	CPSC	Benton 4-in-1 Convertible Crib
Deep Fry Pan	H-E-B Kitchen & Table 5.5 qt. Sauté Pans	CPSC	Cooks Standard NC-00346 5-Quart
Bib	MATVRÅ Infant Bibs	CPSC	Tommee Tippee Closer to Nature
Learning Walker	Crate and Barrel Activity Push Walkers	CPSC	VTech Sit-to-Stand Learning Walker
Heat Pumps	Carrier ductless heat pumps	CPSC	SPLIT INVERTER INDURAMA 12000 BTU
Hair Drier	Xtava Allure/Allure Pro	CPSC	Philips Thermoprotect Hp 8230/00 2100w
Flex Lamp	West Elm Table Lamps	CPSC	LED TaoTronics Flexo
Gas Cooker	Active & Co Portable Single Burner	Product Safety Australia	CAGO JV 02 Gaskocher

Appendix C. Consent Form.

Consent Form for *Human Factors and Engineering Psychology (MSc)*

YOU WILL BE GIVEN A COPY OF THIS INFORMED CONSENT FORM

Please tick the appropriate boxes **Yes No**

Taking part in the study

I have read and understood the study information dated 13/12/2019, or it has been read to me. I have been able to ask questions about the study and my questions have been answered to my satisfaction. ○ ○

I consent voluntarily to be a participant in this study and understand that I can refuse to answer questions and I can withdraw from the study at any time, without having to give a reason. ○ ○

I understand that taking part in the study involves performing a memory task and filling in demographic information such as age, gender, or field of study, as well as filling in 3 personality scales (Geekism, Technology Acceptance and Trust Score). ○ ○

Use of the information in the study

I understand that information I provide will be used for latter data analysis, to measure memory recognition, as a part of a master's degree thesis. ○ ○

I understand that personal information collected about me that can identify me, such as [e.g., my name or where I live], will not be shared beyond the study team. ○ ○

Future use and reuse of the information by others

I give permission for the demographic and personality data, as well as the results corresponding to my performance in the memory experiment, that I provide to be archived in researcher's database so it can be used for future research and learning. ○ ○

Signatures

______________________ ______________________ ________

Name of participant [printed]

Signature Date

I have accurately read out the information sheet to the potential participant and, to the best of my ability, ensured that the participant understands to what they are freely consenting.

_________________________ ___________________ ________

Researcher name [printed] Signature Date

Study contact details for further information: ***Jaume Agud Morera, (0034)-646.774.708***

Contact Information for Questions about Your Rights as a Research Participant

If you have questions about your rights as a research participant, or wish to obtain information, ask questions, or discuss any concerns about this study with someone other than the researcher(s), please contact the Secretary of the Ethics Committee of the Faculty of Behavioural, Management and Social Sciences at the University of Twente by ethicscommittee-bms@utwente.nl

Appendix D. Protocol.

The study sample is going to be composed by a total of 30 participants.

Each participant is invited individually to do the experiment, choosing a silent and comfortable space to work at, and where we will install the computer to use for the experiment. As a general rule, 16” screen monitors will be used, and any change in the size will be reported as this might affect the performance.

Before starting the experiment, the researcher will explain to the participants what is the task going to be like, approximate times, rules, and the fact that his/her performance is going to be assessed using the data obtained during the experiment. This will be done using a participant information sheet. Other questions that the participant might have will be also solved at this stage, although the participants are going to be unaware of the purpose of the thesis in order to avoid affecting their performance. The goal of the experiment will be explained at its end.

After the participant has been told the details about the experiment, a consent form will be provided by the researcher, upon approval by the tutor. Once this is signed, the participant will receive a survey in which they will be asked to fill some demographic and personality data, filling that questionnaire online using the Qualtrics platform.

The researcher will then present the experiment to the participant, using the PsychoPy program on the computer. He/she will be instructed to read at the pre-experiment welcome screen, and asked whether all is clear. The researcher will stress the fact that the keyboard can not be used yet during the presentation of the stimuli. After the participant has seen all the images of the flags (Pre-Experiment), a pre-experiment will be run in order to get a baseline for the memory performance of each participant, provided the participant understood the instructions.

After the pre-experiment phase concludes, the participant will immediately start with the experimental learning phase. Once the participant has been exposed to the 60 pictures of stimuli, he/she will be granted a break of 30 minutes. What to do during that break depends on the participant and location, but as a rule, no exposition to other images or tests should happen during that time, and preferably they are to stay as close to the room as possible.

After the break, the participant will do the final phase, in which he/she has to say whether an image has been shown before by pressing “y” if affirmative, and “n” if negative.

After the test has been done, the participant will be thanked and asked for feedback. After that, the purpose of the research will be revealed, and, if this applies, any reward that could have been discussed per participation (Such as credits or monetary compensation) will be given.

Appendix E. Participant Information Sheet.

You are going to take part in a memory experiment which is part of a master's degree thesis at the University of Twente. The experiment will be performed individually using a 16' computer laptop.

Introduction.

The experiment that you are going to perform aims at exploring your memory performance for images of different stimuli. The researcher of this study is Jaume Agud Morera, student of the Master Degree in Human Factors and Engineering Psychology at the university of Twente (j.agudmorera@student.utwente.nl) supervised by Dr. Simone Borsci (s.borsci@utwente.com), professor of that program.

Purpose of the study.

The purpose of the study is to understand the memory performance of each participant for the images shown, as a part of a study on recognition of stimuli. In this case, performing an experiment is a partial requirement for obtaining the master degree of "Human Factors and Engineering Psychology" at the University of Twente.

Duration and procedure.

The study can be divided in 5 main parts (Questionnaire, Pre-Test, Learning Phase for Experiment, Break, Recognition Task for the experiment). In total, it can take an approximate time of 45 to 60 minutes, including the time of the break (30 minutes).

You may decide how to spend the break, but in order to ensure a good performance in the recognition task, we call for you to not engage in any activity that might involve a cognitive workload, either memory task or attention task.

Rights of participants.

You can quit the experiment at any moment, your data will then be excluded from the database and your previously filled demographic and personality questionnaire will be destroyed.

The usage of the data gathered with the questionnaire, as well as the data regarding your memory performance, will be limited to the scope of this study, and will compose (Anonymously) the final dataset regarding the average performance of all the participants, which will appear in the final thesis report publicly available.

Contacts.

In case of any questions, you can contact the researcher by email (j.agudmorera@student.utwente.nl) as well as by phone (+34-646.774.708).