Social Media & Privacy Issues

Rake, F. E. January 2017

Is there a difference in the disclosure of personal information between the 'real' world and the virtual world and what could be explanations for this phenomenon?

BSc Thesis Marketing and Consumer Behaviour (MCB)

Fleur Rake (960830679120)

Mentor: Dr. Ir. A.R.H. Fischer

Course code: YSS-82312

Abstract

It is generally known that information can be put on the social media without overthinking privacy issues, because it is an abstract and anonymous virtual world. Is there more information disclosed on social media than in the real world? A 2x2 between subject experiment was set up with two variables each consisting of two conditions. A sample of 98 students completed the designed survey with questions high or low in risk of personal information, which was spread in two ways to meet the online world and real world simulations. A factorial analyses of variances was conducted to see if the assumption of the difference in disclosure between the real and virtual world was correct. A factorial ANOVA gave insight in the self-disclosure of information high or low in personal information in the virtual world and in the real world. Findings show that students are relatively open in disclosing personal information in the virtual world, as what was expected, but also in the real world, which contained more disclosure than expected. The Privacy Paradox may not be limited to social media, but might also play a role in the real world.

Table of Contents

Situation and problem

When a serious crime is reported on the news or in a newspaper, it is made sure that the privacy of the criminal is ensured. One way to do this is by only mentioning the first name of the criminal or by placing a censor bar on a picture of him or her. If such a crime has been filmed or photographed and is put on the internet by an individual, often is seen that the privacy of the criminal is not taken into account. The video or photo will be placed on social media without any censoring and it might reach a big audience.

When a general practitioner or a pharmacy wants to share its patients' medical files with other caregivers, it is made sure that the patient agrees and gives permission. There have been many discussions about how well the privacy of medical records is ensured and strict rules are set. People do not want their medical records to be shared with 'unknown' people. However, when social media users need to go to the doctor or hospital, or if they feel ill, it is likely for them to post this on the internet and share this with their virtual friends.

Despite the fact that those situations seem very different, they have a contradictory situation in common. In both situations, the privacy in the real world is ensured with strict rules. People seem to value their privacy and the privacy of others. In the virtual world, people do not seem to care about their privacy as much as they say and do in the real world. What makes people differ in their attitude towards privacy in the real world compared to their behaviour in the virtual world?

First of all, it is important to understand the meaning of a virtual world. A virtual world is a computer-based community where people can interact in a simulated world. Virtual worlds contain settings in which users are not hindered by social barriers or physical constraints of the real world (Lomanowska & Guitton, 2014). Famous virtual worlds are 'Second Life' and 'World of Warcraft'. These are examples of online worlds where avatars (virtual characters) are created by users that interact with other avatars. Social media are a form of virtual worlds too. Avatars do not play a role in social media, but they are still a virtual platform where people can connect and interact with each other and are not hindered by some real life barriers or constraints. Big players in the world of social media nowadays are Facebook, Twitter and WhatsApp (Chaffey, 2016).

Social media is a relatively new concept. The era of social media started with the first form of a social network, called 'Open Diary', which was established by Bruce and Susan Abelson in 1997. Around the same time high-speed internet access started to emerge. Due to high speed internet, the popularity of social media grew. Social platforms like 'Facebook' and 'MySpace' were created and started growing. Facebook, Twitter, Instagram, YouTube,

Whatsapp, gaming websites and many more platforms are helping us to communicate with each other very easily. Nowadays, using social media has become a way of life for the young generation (Chowdhury, 2016).

Relevance

According to the statistics of January 2017, the world counts a total population of 7.476 billion people (Worldometers, 2017). At the same time, the number of active social media users was estimated at 2.307 billion people (Chaffey, 2016). Those statistics show us that about one-third of the world population uses social media actively and all those users have to deal with privacy issues on social media.

Social media is a relatively new topic. This means that some aspects around this topic have not been fully researched yet and there are still topics left to be investigated. This also means that social media users do not have all the knowledge or experiences to know what happens with their online personal data. This lack might lead to users disclosing more personal information than intended. There are many aspects concerning social media, users have no idea of. An example of an action done with the personal data of social media users, is that businesses place 'trackers' on sites and online platforms to track the online behaviour of the user. Those businesses gather information and are aiming to form a profile for every social media user. A data dealer located in Amsterdam, the Netherlands, possesses per Dutch household, with a total of around 7.4 million households in total, more than seventy features. Those features contain for example information about wealth, family formations and buying behaviour (Goslinga, 2015). Data dealers can sell those profiles to interested companies who can use this for online marketing purposes such as online advertisements (VARA: Zembla - Data: Het nieuwe goud, 2015). This might sound as a violation of privacy, but is not illegal. With the use of social media, users might not be conscious about who has access to your files and what happens with them. All social media users agreed to this. How is it possible that social media users are unaware about this?

The conditions of using social media are in detail described in the end-user license agreements (EULA's), including all the information about what will be done to personal information. The user has to agree to the conditions in order to use the platform. The problem is that people often do not read those EULA's. Research shows that only one or two out of every thousand social media users chooses to access the license agreements and that those people mostly do not read for more than one or two seconds, while EULA's often exist of pages full with terms and conditions (Bakos, Marotta-Wurgler, & Trossen, 2009). This shows that there is only a handful of people that read and understand the EULA's they agree to for using social media.

Behaviour towards privacy seems to differ between the ‘real’ world and the virtual world. The virtual world is relatively new and it seems that people are not well informed about what can be done with their online personal information. The goal of this research is to find out whether there is a difference between the disclosure of personal information in the ‘real’ world compared to the virtual world, and how this difference could be explained. This leads to the following research question:

Is there a difference in the disclosure of personal information between the ‘real’ world and the virtual world and what could be explanations for this phenomenon?

Theoretical framework

The two main concepts in this research are 'privacy' and 'the virtual world'.

Privacy

There have been a lot of privacy studies about consumer willingness to provide information (e.g. Bart et al. 2005; Schoenbachler and Gordon 2002), but until now, only a small number of studies have focussed on the degree to which these intentions might influence the behaviour and if there are other factors that might affect this relationship. Risk and trust are two factors that have been investigated regarding privacy intentions and concerns (Bart et al. 2005; Hoffman, Novak, and Peralta 1999; Horne and Horne 2002; Schoenbachler and Gordon 2002; White 2004), but there has not been payed much attention to actual privacy behaviours. It is important to understand how risk and trust might have an influence on the intention to disclose and to the actual disclosure behaviour (Norberg, Horne, & Horne, 2007).

Risk

Risk can be defined as the concern for a negative outcome (Havlena & DeSarbo, 1991). The perceived likelihood that a negative event will occur and the perceived harshness of that specific event influences someone's evaluation of risk (Peter & Tarpey, 1975). In general, the negative outcome of a risk is perceived greater than the potential benefit, for example in the disclosure of information (White, 2004). The negative perception can affect individuals physically, materially or emotionally (Moon, 2000).

Trust

Previous studies have defined 'trust' in a variety of ways. According to Moorman, Deshpande and Zaltman (1993), trust is a willingness for relying on exchange partners. Trust operates differently in online and offline environments. While noting this, Bart et al. (2005) developed a list of factors that have an impact on online trust. Research of both Schoenbachler and Gordon (2002) and Hoffman, Novak and Peralta (1999) found a positive relation between the level of trust and the willingness to provide personal information. The higher the levels of trust, the higher the willingness to provide personal information.

Milne and Boza (1999) suggest that behaviour is directly influenced by trust. Risk considerations also influence the intention to disclosure, but those considerations seem not strong enough to influence the behaviour (Norberg, Horne, & Horne, 2007).

The conceptual model from Norberg, Horne and Horne (2007) suggests that there is no connection between the intention and the actual behaviour and shows how trust and risk operate with regarding to actual behaviour and the behavioural intention (Figure 1). This is in

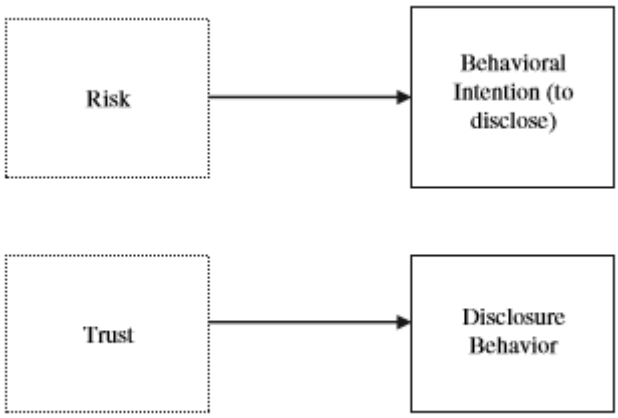
contrast to earlier research of Icek Ajzen and his 'Theory of Planned Behaviour, which states that behavioural intention is the proximal determinant of actual behaviour (Ajzen, 1991).

Norberg, Horne and Horne (2007) argue that in the virtual world, the behavioural intention is not a predictor of actual behaviour because risk influences the intention to disclosure, but there is a trust heuristic which operates in actual disclosure behaviour.

So, in the real world, according to the 'Theory of Planned Behaviour', there is a relation between the behavioural intention and the actual behaviour, while Norberg, Horne and Horne argue that there is no relation.

FIGURE 1

Conceptual Model – Privacy Paradox



The difference between the behavioural intentions to disclosure personal information and the actual disclosure of personal information behaviour is called 'the Privacy Paradox'. Individuals seem to be concerned about their privacy on the Social Web, but this is not reflected in their behaviour. Findings in the research of Norberg, Horne and Horne (2007) support their argument that there is a big gap between the behavioural intentions and the actual behaviour of individuals. Individuals provide significantly more personal information on the internet than what they say they will. This 'Privacy Paradox' leads to the first hypothesis in this research:

Hypothesis 1: '*In the virtual world, more personal information is disclosed than in the real world.'*

The research of Norberg, Horne and Horne (2007) gave insight in 'the Privacy Paradox', but has not fully explained the paradox.
It is assumed that the major reason for 'the Privacy Paradox' is a lack of risk awareness and the users' lack of awareness of possibilities for their privacy protection in the virtual world (Acquisti & Gross, 2006; Boyd & Hargittai, 2010; Debatin et. al., 2009; Tufekci, 2008). Also is supposed that there is an underestimation of the privacy dangers of self-disclosure on the web. This lack of knowledge is partially explained by the fact that most social media users do not, or barely read the end-user license agreements (Bakos, Marotta-Wurgler, & Trossen, 2009).

According to Wheeless and Grotz (1976), self-disclosure is the provision of personal information. It includes 'any message about the self that a person communicates to another'. Self-disclosure can vary a lot between intimacy, value or depth. Self-disclosure is the result of considering utility and risk (Petronio, 2002).

Aspects like gender, cultural background, level of activity and internet experience all have been researched. Monika Taddicken (2013) researched the impact of personality and the relevance of perceptions of the Social Web. Her aim was to find out how much personal information and what kind of information users disclose in the Social Web. The amount of self-disclosure is often formed by the tension between on the one hand the desire to self-disclose, and on the other hand the desire to protect privacy.

In both the real world and the virtual world, it might be that it is not hard for people to disclose insensitive and impersonal information. The tension between the desire to self-disclosure and the desire to protect privacy is low, because there is no need to protect privacy due to the insensitive and impersonal information. Examples of insensitive and impersonal information are date of birth, place of birth and current residence. The more personal and sensitive the information becomes, the more likely it is for people to start considering the benefits of disclosing this information and the disadvantage on their privacy. Examples of more sensitive and more personal information are medial ailments and income. Personal and sensitive information can be named as information with a high risk. The difference between the disclosure of high and low risk information leads to the following hypothesis:

Hypothesis 2: 'Potentially high risk personal information is less likely to be disclosed than low risk personal information.'

A definition for privacy is the right of individuals to determine for themselves when which information is communicated to others and when and how this happens (Westin, 1967).

Altman defines privacy as “selective control of access to the self”. Privacy should not be perceived as something that can reach a maximum. Privacy is a dynamic process of intervention between disclosure and retreat. The ideal point of privacy is reached when there is a balance between the individual need for social interaction and the need for privacy (Altman, 1975).

Virtual world

A big part of the virtual world is covered by social media. Andreas Kaplan and Michael Haenlein define social media as “a group of Internet-based applications that build on the ideological and technological foundations of Web 2.0, which allows the creation and exchange of user-generated content” (Kaplan & Haenlein, 2010). Social media provide value by letting users make content and subscribe to the content of others.

An effective way of reaching as many people as possible which is used more and more, is through social media. Almost one-third of our world population uses social media actively (Chaffey, 2016), so there is a big chance by reaching a large target group by communication through social media. Social media can be used for safety messages, weather alerts or evacuation instructions. Many public safety organizations worldwide are using different social media for spreading their messages, instructions or warnings (Werner, 2011). Companies also communicate and mainly advertise through social media. Through social media and with the use of the personal profiles businesses make of online behaviour and information, it can be perceived as easier to reach a specific target group. The year 2015 was the first year where the European expenditure on online advertisements became higher than the expenditures on TV-advertisements with more than three billion euro (Adex Benchmark, 2015).

Social media has become a great marketing tool, because it is a different way of connecting with the community (Werner, 2011). As mentioned before, there are businesses that collect personal information through social media. They make a profile of every social media users and sell those to businesses that can use the profiles for targeted and personal online advertising or other purposes (VARA: Zembla - Data: Het nieuwe goud, 2015).

A major difference between the real world and the virtual world is ‘the Privacy Paradox’. People seem to be concerned about their online privacy, which is shown in their behavioural intentions, but this is not reflected in their actual behaviour. In the real world there is a link between intentions and behaviour (Ajzen, 1991).

The conceptual model of Norberg, Horne and Horne (Figure 1), shows that there are two factors that influence people's intentions and their actual behaviour. Behavioural intention is influences by risk, and disclosure behaviour is influenced by trust.

According to earlier researches, it is expected that no difference is seen in the disclosure of low risk information between the virtual world and the real world. There is an expectation that the disclosure of high risk information between the virtual world and the real world do differ. The disclosure of high risk information is expected to be higher in the virtual world than in the real world. This interaction between the first and second hypotheses leads to the third and last hypothesis in this research:

H*ypothesis 3: In the virtual world, the difference in probability of providing high and low risk is smaller than in the real world.*

Method

Design and manipulation

A quantitative 2x2 between subject experiment was conducted with two manipulated factors with each two each. The amount of disclosure of personal information was tested against the world it was disclosed in.

The first condition was the world where the participants were assigned to. There were two worlds, the virtual world and the real world. The virtual world was simulated with an online survey, which could be completed on any computer, laptop, tablets or smartphone. Personal disclosure on social media can also be done on computers, laptops, tablet or smartphones. The online survey simulated online conversations and information-sharing with unknown or barely familiar virtual "friends". The real world was simulated by conducting the survey in a face-to-face situation. This simulated a real conversation in daily life with unknown people.

The second condition that was measured is risk. The difference between a high risk and a low risk was analysed. The high risk condition was simulated with personal and privacy-sensitive questions like someone's uncertainties and sexual lifestyle. The low risk condition was simulated with questions that were less personal and low in privacy-sensitivity. Questions like the supermarket someone goes to for groceries, or the number of siblings someone has were asked.

The following table gives an overview of the four conditions in this experiment (Table A).

TABLE A

Design 2 x 2 experiment – between subject design

	Virtual world	*Real world*
	Group 1	Group 2
Low risk	*Online survey with questions low in privacy-sensitivity*	*Face-to-face survey with questions low in privacy-sensitivity*
	Group 3	Group 4
High risk	*Online survey with questions high in privacy-sensitivity*	*Face-to-face survey with questions high in privacy-sensitivity*

A survey was composed and a pilot face-to-face survey was conducted to test whether the high risk questions were too personal and sensitive, so that people would not want to answer them. It turned out that the questionnaire did not contain questions that were privacy-sensitive enough. An example of a question was: 'What is your sexual preference?'. It turned out that people did not have doubts about answering the question. The explanation for this could be that heterosexual people have no difficulty answering this question, because heterosexuality is seen as 'normal'. Only homosexual or bisexual people who are also not open about it, might have difficulty answering this question. In general, the questions that were made up at first were only 'hard' to answer for people with an exceptional answer. A small research was done by asking people what questions would be hard for them to answer. A new survey was composed with high risk questions that were possibly too personal for almost everyone to answer.

Students are likely to give answer to a survey sooner if it is conducted in name of the Wageningen University. The reason for this is that people might want to participate to surveys sooner if it is for scientific research. For this reason, the GMOI (Gelders Marktonderzoek Instituut), a market research company, was made up as the company that conducted the interview. This company was made up to make sure people would answer the survey as they would do if a normal 'stranger' would ask them the questions.

In the debriefing at the end of the survey, the respondents were informed about the fact that the GMOI does not exist. An explanation was given why that company was made up and that the survey was actually conducted for a scientific research. An e-mail address was given for any suggestions or questions about the research (see Appendix I for the final survey).

Measures

The survey consisted of three parts. The first part consisted of demographical questions. Four questions were asked about the year of birth, gender, main daily activity and the highest accomplished education level. A randomisation check was done to see if the groups formed by manipulation were equal.

The second part of the survey consisted of the actual manipulation. For both high and low risk, seven questions were asked. To measure the difference between the two groups, a scale was made. The answers in particular did not matter, but the number of questions answered did matter. So, if the value was zero, this meant that the participant gave 'no answer' to the questions seven times, so the participant gave zero times any answer. A value of seven means that the participant answered all the high or low risk questions, so the

participant never gave the answer ‘no answer’. This meant that the person did not have any trouble with answering the personal questions.

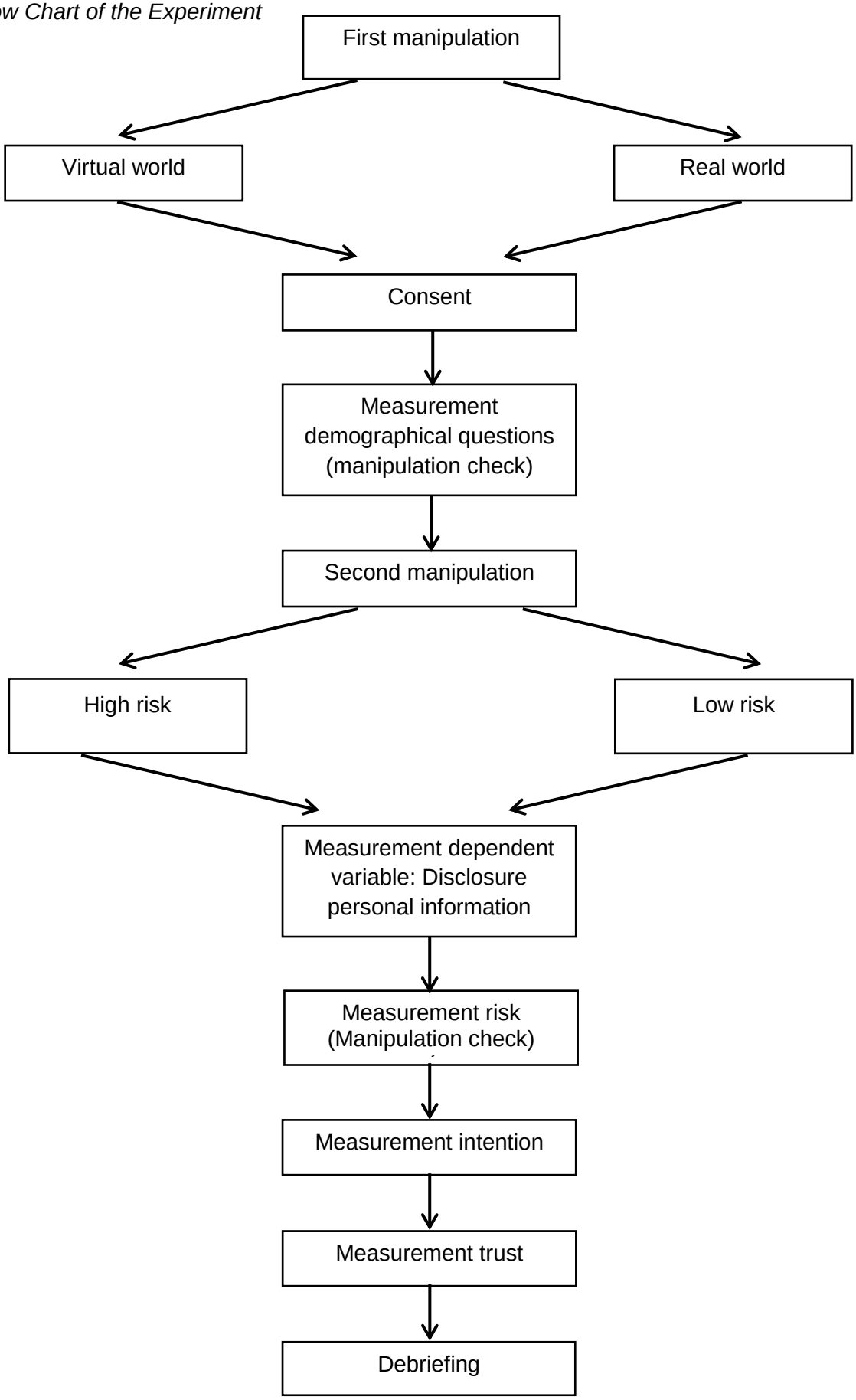
The last part of the survey consisted of control questions and manipulation checks. Questions were asked to see how personal the questions were experienced by the participants. This question would point out if the manipulation worked out and if the reason for not answering the questions was a privacy-issue, or if there was another reason. A Likert scale was used with the following options: not personal at all (1), a bit personal (2), neutral (3), pretty personal (4) and too personal (5). If a participant has the value zero, which means that the participant did not answer any high or low risk questions, it is assumed that this is due to the level of the privacy-sensitivity of the questions. This manipulation question is added to the survey to see if the privacy-sensitivity was successful, and if that privacy issue was the reason for not answering the questions.

A following question asked if the participant would give the same answers in future researches to other companies. This question measured the intention. Previous research said that mainly in the condition with online high risk personal questions, the intention of people is to not disclose that much information, but their behaviour shows that they disclose more information than that. The measurement of intention has been involved in the manipulation check.

The last questions asked if the participants trusted the GMOI. It is important to know if the participants trusted the GMOI or the person conducting the survey in name of the GMOI. The reason is that if the participants did not trust it, it is not likely for the participant to answer the questions. Trust had to be the same in every condition, because it could not have influence the amount of disclosure. See Figure 2 for the flowchart of the experiment.

FIGURE 2

Flow Chart of the Experiment



Participants

There are different norms and values for privacy for every country or culture in the world. This makes it hard to compare privacy across countries. To make this research internally more valid, this research was focused on one culture and one understanding of privacy. The experiment was conducted in the Netherlands and was therefore written in Dutch. Expected was that the different conditions would not be different in the Netherlands than in other Western countries.

The biggest group using social media is the group of youngsters and young adults from 15 to 29 years old (Van der Veer, Sival, & Van der Meer, 2016), but also older people (55 years and older) start using social media more and more (Madden, 2010). It was not possible to ask youngsters under 16 due to the required permission from their parents or guardians. Therefore, the experiment was conducted by people from the age of 16 and older. For this research, it was most useful to conduct the survey among the most frequent social media users. Therefore, the survey was conducted mostly among students of all levels of education. For the simulation of the virtual world (the online survey), people needed to have access to the internet.

Procedure

When a person participates in this experiment, he or she assumed that their participation is for the questionnaire of a market research agency (GMOI). A person participated this experiment either by being addressed to in the face-to-face experiment, or by the online survey. This manipulation decided whether the participant was located in the virtual world or in the real world. There had to be made sure that a participant did not participate in two or more conditions. To prevent this, the respondent was asked before participating if he or she already completed the other form of the survey (online or face-to-face). This was mentioned in the introduction of the survey.

Online survey

The online survey was spread and shared through Facebook. Expected was that this would reach a large range and variety of people who fitted in the target group. The online survey was conducted through Qualtrics. This program makes sure people will be randomly manipulated. It also made sure that both conditions occur almost evenly. The option that every question must be answered was used. This option was chosen to be sure that people have considered answering ‘no answer’. If the forced response was not used, reasons for not answering a question could also be because people were in a rush or accidentally left the question open.

Face-to-face survey

The face-to-face survey was conducted at the Wageningen University. This seemed the best place to meet a lot of students and PhD students. The survey was conducted during the lunch break, to reach the greatest variety and amount of people. A pile of printed surveys was used for conducting the survey. Before the conduction, the high risk and low risk surveys were randomly put together to make sure the manipulation was as random as possible. One goal was to ask as many different people that fitted in the target population. Therefore, random males and females of different ages were asked. To make sure the survey was really a face-to-face survey, the person conducting the interview asked the questions herself. The reason is that if people would get the survey on paper and fill it in themselves, this might have been similar to the online environment, because it would still be impersonal and might feel anonymous.

The amount of participants needed for this research was set at eighty. This seemed a good amount, because this meant that there will be forty respondents in each world. This lead to at least twenty participants in each condition.

Data analysis

The analyses of the data were performed with the help of the computer program SPSS.

This experiment contained two independent variables (world and risk) with each two conditions (virtual vs. real world and high vs. low risk) analysed with a factorial analysis of variance. In general, a factorial ANOVA test is done to see if there is an interaction between the two independent variables on the dependent variable. So, the ANOVA test would tell if there is a correlation between the type of world and the amount of risk on the disclosure of information. The main effect for the condition of the two different worlds, the main effect for the condition of the two amounts of risk and the interaction effect of these two variables or conditions have been tested.

A randomisation- and manipulation check were done. Extra variables were made before starting the analyses to merge data about type of world, amount of risk and the total disclosure. After adding the extra variables, several analyses were conducted with the assumption of finding any significant results.

Results

The dependent variable was measured at a ratio level scale with eight levels (a disclosure score of zero to seven). The two independent variables consisted each of two categorical independent groups. There was an independence of observations because each condition consisted of unique participants.

Randomisation

All conditions contained more than twenty respondents. The condition in the virtual world with low risk personal information contained the most respondents (Appendix II, Table 1). Of all 98 respondents, in total almost two times as much females than males particpated. In all four conditions, more females than males participated.

In both conditions, the average year of birth was close to 1995 with relatively small standard deviations (Real world: SD = 2.10, Virtual World: SD = 4.01). The single respondent with 1969 as year of birth did not have influence on further data analyses.

The main daily activity of the plurality of the respondents is studying. Only seven respondents have another main daily activity. (Appendix II, Table 3). The main highest completed level of education was in both conditions VWO and BSc (Appendix II, Table 4).

In total, 89 respondents answered all seven questions. Nine respondents deviated and answered six out of seven questions.

Factorial ANOVA

Levene's test was statistically significant, so there is no equality in the variances, $F(3, 94) = 9.435$, $p = 0.00$.

The Descriptive Statistics shows relatively small deviations. The homogeneity of variances was not violated due to the small standard deviation. However, the standard deviation of low risk personal information questions in the real world was zero, because all the questions in this condition were answered (Table B).

TABLE B

Mean and Standard Deviation of Disclosure per Condition

	Risk	Mean	Std. Deviation
Virtual world	*High*	6.870	0.344
	Low	6.943	0.236
	Total	6.914	0.283
Real world	*High*	6.800	0.410
	Low	7.000	0.000
	Total	6.900	0.304

The mean of the total disclosure in the virtual world was comparable to the total disclosure in the real world (Table B). The effect of the type of world was not statistically significant (Table C).

The mean of the disclosure in the high risk condition was lower than the mean of the disclosure within low risk condition (Table B). The Tests of Between-Subject Effects show that the effect of the amount of risk was statistically significant (Table C). High risk personal information leads to less disclosure.

The interaction effect shows no statistically significant effect (Table C). An additional simple effect analyses was conducted for the interaction effect. Outcomes suggest that there was a statistically significant effect in the real world within the amount of risk, $F(1, 94) = 4.89$, $p = 0.029$, $\eta^2 = 0.05$, 95% CI (0.02 – 0.38). The lower bound and upper bound were both above zero.

TABLE C

Effect of independent variables and interaction effect

Effect on Disclosure	df	F	p	η^2
Type of world	1, 94	0.01	0.917	0.05
Amount of risk	1, 94	5.30	0.023	0.05
Interaction effect	1, 94	1.14	0.288	0.05

Additional analyses

Most unanswered questions appeared mostly in the high risk condition (Table B, see lowest mean). An independent T-test showed that in the real world people gave a higher average for the amount of having sex per month than the virtual world. The amount of grocery shopping per week was in the real world averagely 0.7 times of doing groceries higher than in the virtual world (Table D).

TABLE D

Average, t-value and p-value of unanswered questions

	Real world	Virtual world	t	p
Average number of sex	7.3	4.9	t (34) = 1.72	0.212
Average number of groceries	4.2	3.5	t (47) = 1.31	0.196

Manipulation

A cross-tab showed that in the low risk survey, almost every participant had the intention to give the answers on the questions for other future researches or companies. In the real world and the virtual world both, most of the participants had the intention to give the same answers to other future researches or companies. A significant difference was found in the intention between the high risk condition and the low risk condition ($t(96) = 3.58$, $p = 0.001$) (Appendix II, Table 5).

As expected, the difference between the sensitivity of the questions in the high risk ($M = 3.47$, $SD = 1.14$) and the low risk questions ($M = 2.09$, $SD = 0.91$) was statistically significant ($t(96) = 6.64$, $p = 0.00$). An additional effect was found between the sensitivity of the questions the real world and the virtual world ($t(96) = -2.095$, $p = 0.039$). The questions in the virtual world were experienced more sensitive than in the real world.

Analyses showed that there were no statistically significant effects for the trust in GMOI for the different variables. It was important that the trust was almost the same in every condition, so it had no influence on the amount of disclosure. (Appendix II, Table 6).

Discussion

The first hypothesis of this research expected the amount of disclosure in the virtual world to be higher than in the real world. The findings show that there is no statistically significant difference in the amount of disclosure between the real world and the virtual world. A reason might be that there is no actual difference or it might be the case that the simulation of the two types of world might have not worked as it should. This simulation might have had a smaller or limited effect on the amount of disclosure. A possible solution might be sharing information with a stranger in the real world versus the virtual world. Would creating a fake account with a person asking you all kinds of personal information on a social medium lead to more disclosure than in a real conversation with a stranger? Future research could use this perspective in the experiment in order to simulate the two types of world better.

According to the second hypotheses, expected was that the high risk personal information was less likely to be disclosed than the low risk personal information. The analyses conducted can conclude that a statistically significant effect is found in the disclosure between the high risk personal information and the low risk personal information. This is a logical finding, but the effect could have been greater. The choice for conducting an ethically correct survey has a large influence on this effect. To stay within the borders of ethicality, the high risk personal questions are not as personal as they could be. This demarcation was a conscious choice for conducting an ethically correct experiment. There is no regret in this choice. Asking unethically information might be too sensitive or rude, which can cause participants to react angry or sad about possible past experiences. These reactions are prevented, because they are not necessary and they do not add anything to the experiment.

One of the six ‘Weapons of Influence’ of Cialdini (2009) might have caused the effect to not be as large as it could be. The weapon ‘liking’ means feeling similar with a person and therefore feeling more comfortable in doing the desired behaviour. This means that the respondents possibly have felt similar with the person conducting the survey. It is likely that this effect had a greater influence on the face-to-face surveys. Having a regular, not outstanding student conducting the surveys with students as respondents might have led to more disclosure of personal information. It is generally known that students are more open about their personal information, even more when also a similar person like them conducts the survey. Future research could make use of a person conducting the interviews who is not similar with the respondents. Expected is that this will lead to less disclosure. Because the group of non-similar people than the particular respondent is expected to be relatively larger than the group with similar people, using a non-similar person as conductor of the surveys will lead to the most real simulation of the situation.

Another theory of Cialdini called ‘Foot in the door’ has had an influence in the way that people accepted to participate the survey. Human beings prefer having consistent behaviours and attitudes (Cialdini, 2009). By accepting and participating in the survey, it is inconsistent to choose not to answer a question. The commitment to participate in the survey was already made and will probably not be changed. To avoid inconsistent behaviour, participants might have felt they had to answer questions, despite the fact they do not always have to feel comfortable with this choice.

Another reason for the respondents answering the questions is the context they were in. The person conducting the interviews approached people in their lunch time. This made the respondents having the time to answer, not being in a hurry or having a deadline. This may have played a role in a more positive vibe leading to a larger amount of self-disclosure. It was made sure that the people approached were alone and not in peer-groups. This led to less fear of the opinions of the peer-groups of the respondents and no social pressure. This might have had a more relax setting leading to more disclosure.

In general, it was noticeable that students wanted to contribute to the survey because they felt like contributing to scientific research, despite the fact that this information was not given beforehand. The word ‘research’, in their own words, triggered students to contribute to the survey. Displaying the research like a fully commercial research might lead to different findings, like students daring to say they do not want to participate or students daring to say they do not want to answer a question.

The third and last hypothesis tested whether there is a statistically significant effect found in the interaction of the two independent variables. Unfortunately, the interaction between the two independent variables does not show a statistically significant effect. This was a reason to conduct an additional analysis. A simple effect analyses suggests that there is a statistically significant effect in the amount of disclosure the real world between the high risk personal information and low risk personal information. One criticism is that the homogeneity of variances is violated, which requires that the findings are interpreted with caution.

Since outcomes were not as expected according to the literature, additional analyses were conducted to see if there is any effect found in the content of the unanswered questions. All unanswered questions fell into the same category, the open questions. There are no significant differences found between the unanswered questions in the real and virtual world, but the difference between the amount of sex in the virtual world compared to the real world is notable. The respondents in the real world score 2.5 times per month higher than in the virtual world. This is not a significant outcome because there is an enormous standard deviation. A possible reason could be that the respondents gave a socially desirable answer,

because sex is a personal topic. It can be the case that they were ashamed to name a low number, because they have the feeling that would be weird. This would probably have played a smaller role in the virtual world, because the online survey is totally anonymous. In our society nowadays, and among students it is experienced as normal to have sex and talk about it since students are known as people with relatively much self-disclosure. By possibly feeling some peer-pressure and to prevent being different than the majority, it is likely that respondents exaggerated and named a higher number in the fear that a lower number is not 'normal'.

The measure used in this research seems to work in this kind of research, bus was not sensitive enough. The sensitivity of this measure needs to be improved for further research in order to determine larger effects in the amount of disclosure within the different conditions. This can be done in two ways. There could be done more research to more sensitive questions within the borders of ethicality, and there could be thought of more questions, which hopefully makes the spreading between the amount of disclosure larger.

The majority of the respondents were students. It is generally known that students are quite open and social on social media, but this research shows that this is also the case in a face-to-face survey. This might have had influence on the outcomes. They might have less difficulty disclosing personal information than other groups in our society. Another think which definitely had an influence on the outcomes of this research is that students grew up with the rise of computers, internet and social media. Because they do not know different and because they are used to the use of social media they might 'dare' to disclose much more information than older groups in our society, for who social media is less familiar and known. In general, students and teenagers have more experience with social media and its consequences. They are not afraid of online disclosure. This will also be the case for younger generations that are also raised with the virtual world close to them.

A negative side to the younger generation having been raised with social media is that things usually happening in daily life may have been replaced to the virtual world. An example is cyberbullying. Cyberbullying became a problem when more and more people started using social media. Not only self-disclosure is easier through social media, but also disclosure of things about others, which happens in a negative way with cyberbullying. It is easier to say negative things to people on the internet and the impact of these negative things might be bigger than the person saying it expects. Why does this happen? Are people being more honest in the virtual world, or are they deliberately saying bad things because it is anonymous and easier? It appears that the virtual world has different norms and conventions towards disclosure. But, what could be a reason that the questions in the virtual world were

experienced more personal than in the real world? This can be an interesting angle of incidence for future research.

This research gave insight in the self-disclosure of information high or low in personal information in the virtual world and in the real world. Findings show that students are relatively open in disclosing personal information in the virtual world, as what was expected, but also in the real world, which was more than expected. The Privacy Paradox may not be limited to social media, but might also play a role in the real world. Future research might investigate and point out when privacy issues actually start playing a role. The no-answer scale for this research has worked for this kind of research, but the sensitivity needs to be improved. This could be done by creating more questions and perhaps by making no-answer the default option.

Reference list

Acquisti, A., & Gross, R. (2006, June). Awareness, information sharing, and privacy on the Facebook. Paper presented at the 6th ‘‘Privacy Enhancing Technologies,’’ workshop Cambridge, England.

Ajzen, A. (1991). The Theory of Planned Behavior. *Organizational Behavior And Human Decision Processes*, *50*(2), 179-211. Doi:10.1016/0749-5978(91)90020-T

Altman, I. (1975). The environment and social behavior: Privacy, personal space, territory, crowding. Monterey: Brooks/Cole Publications.

Bakos, Y., Marotta-Wurgler, F., & Trossen, D. (2009). Does Anyone Read the Fine Print? Testing a Law and Economics Approach to Standard Form Contracts. *SSRN Electronic Journal*. Doi: 10.2139/ssrn.1443256

Bart, Iakov Y., Venkatesh Shankar, Fareena Sultan, and Glen L. Urban. 2005. Are the Drivers and Role of Online Trust the Same for All Web Sites and Consumers? A Large Scale Exploratory Empirical Study. Journal of Marketing. 69 (October): 133–152.

Boogaard, M. & Ytsema, N. (2015). *VARA: Zembla - Data: Het nieuwe goud*. Hilversum, the Netherlands.

Boyd, D., & Hargittai, E. (2010). Facebook privacy settings: Who cares? First Monday,15(8). http://www.uic.edu/htbin/cgiwrap/bin/ojs/index.php/fm/article/view/3086/2589 (June 6th, 2011).

Chaffey, D. (2016). *Global Social Media Statistics Summary 2016*. *Smart Insights*. Retrieved 19 September 2016, from http://www.smartinsights.com/social-media-marketing/social-media-strategy/new-global-social-media-research/

Chowdhury, A. (2016). Social Media Today. The Financial Express

Cialdini, R. (2009). *Influence - Science and Practice* (5th ed.). Boston: Pearson Education.

Debatin, B., Lovejoy, J. P., Horn, A.-K., & Hughes, B. N. (2009). Facebook and online privacy: Attitudes, behaviors, and unintended consequences. Journal of Computer-Mediated Communication, 15(1), 83–108.

Goslinga, M. (2015). |Wie zijn de grootste datahandelaren van Nederland (en hoe beïnvloeden zij ons leven)? *De Correspondent*. Retrieved from https://decorrespondent.nl/3159/Wie-zijn-de-grootste-datahandelaren-van-Nederland-en-hoe-beinvloeden-zij-ons-leven/343753822269-78966ab1

Havlena, W. & DeSarbo, W. (1991). On the Measurement of Perceived Consumer Risk. *Decision Sciences*, *22*(4), 927-939. Doi: 10.1111/j.1540-5915.1991tb00372.x

Hoffman, D., Novak, T., & Peralta, M. (1999). Building consumer trust online. *Communications Of The ACM*, *42*(4), 80-85. Doi: 10.1145/299157.299175

Horne, D. & Horne, D. (2002). Database Marketing: When Does Good Practice Become an Invasion of Privacy?. *Marketing Theory And Applications*, *13*, 480-486.

Kaplan, A. & Haenlein, M. (2010). Users of the world, unite! The challenges and opportunities of Social Media. Business Horizons, 53(1), 59-68. Doi: 10.1016/j.bushor.2009.09.003

Knapp, D., Marouli, E., & Fennah, A. (2015). Adex Benchmark 2015. *IAB Europe*.

Lomanowska, A. & Guitton, M. (2014). My avatar is pregnant! Representation of pregnancy, birth, and maternity in a virtual world. *Computers In Human Behavior*, *31*, 322-331. Doi: 10.1016/j.chb.2013.10.058

Madden, M. (2010). Older adults and social media. Pew Internet & American Life Project, 27. Http://pewinternet.org/Reports/2010/Older-Adults-and-Social-Media.aspx

Milne, G. & Boza, M. (1999). Trust and concern in consumers' perceptions of marketing information management practices. *Journal Of Interactive Marketing*, *13*(1), 5-24. Doi: 10.1002/(sici)1520-6653(199924)13:1<5::aid-dir2>3.0.co;2-9

Moon, Y. (2000). Intimate Exchanges: Using Computers to Elicit Self-Disclosure From Consumers. *Journal of Consumer Research*, *26*(4), 323-339. Doi: 10.1086/209566

Moorman, C., Deshpande, R., & Zaltman, G. (1993). Factors Affecting Trust in Market Research Relationships. *Journal Of Marketing*, *57*(1), 81. Doi: 10.2307/1252059

Norberg, P., Horne, D., & Horne, D. (2007). The Privacy Paradox: Personal Information Disclosure Intentions versus Behaviors. *Journal Of Consumer Affairs*, *41*(1), 100-126. Doi: 10.1111/j.1745-6606.2006.00070.x

Peter, J. & Tarpey, Sr., L. (1975). A Comparative Analysis of Three Consumer Decision Strategies. *Journal of Consumer Research*, *2*(1), 29. Doi: 10.1086/208613

Petronio, S. (2002): Boundaries of privacy. Dialectics of disclosure. Albany, NY: SUNY Press.

Schoenbachler, Denise D. and Geoffrey L. Gordon. 2002. Trust and Customer Willingness to Provide Information in Database-Driven Relationship Marketing. Journal of Interactive Marketing, 16 (Summer): 2–16.

Taddicken, M. (2013). The 'Privacy Paradox' in the Social Web: The Impact of Privacy Concerns, Individual Characteristics, and the Perceived Social Relevance on Different Forms of Self-Disclosure. *Journal of Computer-Mediated Communication*, *19*(2), 248-273. Doi: 10.1111/jcc4.12052

Tufekci, Z. (2008). Can you see me now? Audience and disclosure regulation in online social network sites. Bulletin of Science, Technology & Society, 20–36.

Van der Veer, N., Sival, R., & Van der Meer, I. (2016). Nationale Social Media Onderzoek 2016. *Newcom Research & Consultancy*.

Weber, R. (2015). The digital future – A challenge for privacy? *Computer Law & Security Review*, *31*(2), 234-242. Doi: 10.1016/j.clsr.2015.01.003

Werner, C. (2011). Social Media and Public Safety. *Fire Technology*, *36*(5), 86.

Westin, A. F. (1967). Privacy and freedom. New York: Atheneum.

Wheeless, L. R., & Grotz, J. (1976). Conceptualization and measurement of reported self-disclosure. Human Communication Research, 2(4), 338–346.

White, T. B. (2004). Consumer Disclosure and Disclosure Avoidance: A Motivational Framework. *Journal Of Consumer Psychology*, *14*(1-2), 41-51. Doi: 10.1207/s15327663jcp1401&2_6

Worldometers - Real time wereld statistieken. (2017). *Worldometers.info*. Retrieved 9 January 2017, from http://www.worldometers.info/nl/

Appendices

Appendix I – Surveys

Survey I – High Risk (Dutch)

Gelders Marktonderzoek Instituut

Q5 *GMOI, het Gelders Marktonderzoek Instituut, doet marktonderzoek in opdracht van bedrijven of organisaties. Voor dit onderzoek zijn ze op zoek naar studenten of PhD-ers. De antwoorden die u geeft zullen geheel anoniem blijven. Deze enquête duurt enkel een paar minuten. Sommige vragen kunnen persoonlijk aankomen. U heeft per vraag altijd de optie 'Geen antwoord' als u deze vraag niet wilt beantwoorden. U kunt ook op elk moment stoppen met deze enquête. Het GMOI doet ook via mondelinge enquêtes onderzoek. Als u hier al aan hebt meegedaan wordt u vriendelijk verzocht om niet meer deel te nemen aan deze enquête.*

----- *DEMOGRAFISCHE VRAGEN* -----

Q1 Wat is uw geboortejaar?

_____ Geboortejaar (1)

Q2 Wat is uw geslacht?

- ❍ Man (1)
- ❍ Vrouw (2)

Q3 Wat is uw belangrijkste dagelijkse bezigheid?

- ❍ Studeren (1)
- ❍ Werken (2)
- ❍ Anders, namelijk (3) _____

Q4 Wat is uw hoogst voltooide opleidingsniveau?

- ❍ HAVO (1)
- ❍ VWO (2)
- ❍ MBO (3)
- ❍ HBO (4)
- ❍ Bsc (5)
- ❍ Msc (6)
- ❍ PhD (7)
- ❍ Anders, namelijk (8) _____

----- *VRAGEN RISISCO* -----

Q7 Heeft u zich wel eens slecht of niet tevreden gevoeld over uw lichaam?

- Ja (1)
- Nee (2)
- Geen antwoord (3)

Q8 Bent u onzeker?

- Ja (1)
- Nee (2)
- Geen antwoord (3)

Q9 Heeft u zich wel eens onzeker of buitengesloten gevoeld?

- Ja (1)
- Nee (2)
- Geen antwoord (3)

Q10 Hoe vaak heeft u gemiddeld seks per maand? (Geen antwoord geven is mogelijk)

Q11 Heeft u ooit getwijfeld aan uw seksuele geaardheid?

- Ja (1)
- Nee (2)
- Geen antwoord (3)

Q12 Heeft u ooit een soa-test gedaan?

- Ja (1)
- Nee (2)
- Geen antwoord (3)

Q13 Heeft u ooit seks gehad waarvan u het achteraf liever niet had gedaan?

- Ja (1)
- Nee (2)
- Geen antwoord (3)

----- *CONTROLEVRAGEN* -----

Q14 Hoe persoonlijk vond u de vragen?

- Helemaal niet persoonlijk (1)
- Een beetje persoonlijk (2)
- Neutraal (3)
- Redelijk persoonlijk (4)
- Te persoonlijk (5)

Q15 Zou u de antwoorden op de vragen die u net hebt beantwoord in het vervolg ook aan andere bedrijven geven?

- ❍ Ja (1)
- ❍ Weet ik niet (2)
- ❍ Nee (3)

Q16 Vertrouwt u het Gelders Marktonderzoek Instituut (GMOI)?

- ❍ Ja (1)
- ❍ Weet ik niet (2)
- ❍ Nee (3)

Q17 Bedankt voor het invullen van deze enquête! Ik ben niet helemaal eerlijk geweest. Deze vragenlijst maakt namelijk deel uit van mijn bachelor scriptie. Ik doe onderzoek naar privacy issues op sociale media. Met deze vragenlijst onderzoek ik het verschil tussen de gegevens die mensen van zichzelf blootstellen op sociale media vergeleken met de persoonlijke gegevens die blootgesteld worden in de 'echte' wereld. Het gaat om het aantal vragen dat u heeft beantwoord en niet om de inhoud van het antwoord op de vragen. De antwoorden blijven geheel anoniem en er zal niets mee gedaan worden. Mocht u vragen of opmerkingen hebben over deze enquête? Stuur dan een mail naar fleur.rake@wur.nl

Survey II – Low Risk (Dutch)

Gelders Marktonderzoek Instituut

Q5 *GMOI, het Gelders Marktonderzoek Instituut, doet marktonderzoek in opdracht van bedrijven of organisaties. Voor dit onderzoek zijn ze op zoek naar studenten of PhD-ers. De antwoorden die u geeft zullen geheel anoniem blijven. Deze enquête duurt enkel een paar minuten. Sommige vragen kunnen persoonlijk aankomen. U heeft per vraag altijd de optie 'Geen antwoord' als u deze vraag niet wilt beantwoorden. U kunt ook op elk moment stoppen met deze enquête. Het GMOI doet ook via mondelinge enquêtes onderzoek. Als u hier al aan hebt meegedaan wordt u vriendelijk verzocht om niet meer deel te nemen aan deze enquête.*

----- *DEMOGRAFISCHE VRAGEN* -----

Q1 Wat is uw geboortejaar?

_____ Geboortejaar (1)

Q2 Wat is uw geslacht?

- ❍ Man (1)
- ❍ Vrouw (2)

Q3 Wat is uw belangrijkste dagelijkse bezigheid?

- ❍ Studeren (1)
- ❍ Werken (2)
- ❍ Anders, namelijk (3) _____

Q4 Wat is uw hoogst voltooide opleidingsniveau?

- ❍ HAVO (1)
- ❍ VWO (2)
- ❍ MBO (3)
- ❍ HBO (4)
- ❍ Bsc (5)
- ❍ Msc (6)
- ❍ PhD (7)
- ❍ Anders, namelijk (8) _____

----- *VRAGEN RISICO* -----

Q18 Heeft u een bijbaantje?

- ❍ Ja (1)
- ❍ Nee (2)
- ❍ Geen antwoord (3)

Q19 Heeft u een tussenjaar gehad tussen uw studie?

- ❍ Ja (1)
- ❍ Nee (2)
- ❍ Geen antwoord (3)

Q20 Hoe veel broers en zussen heeft u in totaal? (Geen antwoord geven is mogelijk)

_____ Aantal broers en zussen (1)

Q21 Hoe vaak doet u gemiddeld boodschappen per week? (Geen antwoord geven is mogelijk)

Q22 Heeft u ooit getwijfeld over uw studie?

- ❍ Ja (1)
- ❍ Nee (2)
- ❍ Geen antwoord (3)

Q24 Heeft u ooit gereisd per vliegtuig?

- ❍ Ja (1)
- ❍ Nee (2)
- ❍ Geen antwoord (3)

Q25 Bent u ooit op vakantie geweest waarbij u het achteraf liever niet had gedaan?

- ❍ Ja (1)
- ❍ Nee (2)
- ❍ Geen antwoord (3)

----- *CONTROLE VRAGEN* -----

Q14 Hoe persoonlijk vond u de vragen?

- ❍ Helemaal niet persoonlijk (1)
- ❍ Een beetje persoonlijk (2)
- ❍ Neutraal (3)
- ❍ Redelijk persoonlijk (4)
- ❍ Te persoonlijk (5)

Q15 Zou u de antwoorden op de vragen die u net hebt beantwoord in het vervolg ook aan andere bedrijven geven?

- ❍ Ja (1)
- ❍ Weet ik niet (2)
- ❍ Nee (3)

Q16 Vertrouwt u het Gelders Marktonderzoek Instituut (GMOI)?

- ❍ Ja (1)
- ❍ Weet ik niet (2)
- ❍ Nee (3)

Q17 *Bedankt voor het invullen van deze enquête! Ik ben niet helemaal eerlijk geweest. Deze vragenlijst maakt namelijk deel uit van mijn bachelor scriptie. Ik doe onderzoek naar privacy*

issues op sociale media. Met deze vragenlijst onderzoek ik het verschil tussen de gegevens die mensen van zichzelf blootstellen op sociale media vergeleken met de persoonlijke gegevens die blootgesteld worden in de 'echte' wereld. Het gaat om het aantal vragen dat u heeft beantwoord en niet om de inhoud van het antwoord op de vragen. De antwoorden blijven geheel anoniem en er zal niets mee gedaan worden. Mocht u vragen of opmerkingen hebben over deze enquête? Stuur dan een mail naar fleur.rake@wur.nl

Appendix II – SPSS output

Table 1 – Crosstab type of world and amount of risk (four conditions)

Distribution per condition	*Real world*	*Virtual world*	*Total*
Low risk	20	35	55
High risk	20	23	43
Total	40	58	98

Table 2 – Crosstab males (♂) and females (♀)

Distribution ♂ and ♀	*Real world*	*Virtual world*	*Total*
Low risk	9 ♂, 11 ♀	12 ♂, 23 ♀	21 ♂, 34 ♀
High risk	5 ♂, 15 ♀	8 ♂, 15 ♀	13 ♂, 30 ♀
Total	14 ♂, 26 ♀	20 ♂, 38 ♀	34 ♂, 64 ♀

Table 3 – Distribution Main Daily Activity

Distribution main daily activity	*Real world*	*Virtual world*	*Total*
Study	39	52	91
Work	1	3	4
Other	0	3	3
Total	40	58	98

Table 4 – Distribution Highest Completed Level of Education

Distribution Level of Education	*Real world*		*Virtual world*		*Total*
Risk	*High*	*Low*	*High*	*Low*	
HAVO	0	0	6	6	12
VWO	12	14	11	19	56
MBO	0	0	0	1	1
HBO	3	1	1	2	7
BSc	5	4	5	5	19
MSc	0	1	0	2	3
Total	20	20	23	35	98

Table 5 – Intention

(See Appendix I, Q15 for scale)

Mean of intention per condition	*Real world*	*Virtual world*	*Total*
Low risk	1.25	1.31	1.28
High risk	1.65	2.04	1.85
Total	1.45	1.65	1.55

Table 6 - Trust in GMOI

(See Appendix I, Q16 for the scale)

Mean in trust	*Real world*	*Virtual world*	*Total*
Low risk	1.70	1.63	1.67
High risk	1.30	1.52	1.41
Total	1.50	1.59	1.55