Enhancing privacy and motivation through better survey design

Provision of the taskbar, no-answer-option and sensitive questions

12th of September '13

Erasmus University of Rotterdam

Jeroen Wagenaar

302393jw

Master Marketing, Erasmus School of Economics (ESE)

ERASMUS UNIVERSITEIT ROTTERDAM

Table of content

Acknowledgements

This thesis is written in the period of July 2012 till July 2013. It is the final stage of MSc-education of Marketing at the Erasmus School of Economics (ESE), Erasmus University of Rotterdam.
I have learned to do an individual research project in which a lot challenges are to be face and overcome. Writing this thesis has been an effort for me, both personally and intellectually. Therefore, it has been a very valuable project.

I would like to thank a few people for their support during the realization of my thesis.

Martijn de Jong for his time and effort as my supervisor guiding me while writing, gathering data and conducting the research. Thanks for the time you took to prepare our meetings for over a year and the efforts you made help me looking at research problems a different way.

My girlfriend Mary Bolkenbaas for giving me the psychological strength to work intensively on this thesis while working long hours at Coolblue. Thanks to my parents, who always believed in me and gave me the freedom to finish the way I wanted to finish. Their patience and belief gave me the strength and energy I needed. Also I would like to thank Ingrid Everaars- Oranje as a confidante for being very involved with my struggles and giving me the structure I needed to find the ultimate way in my time management. Saskia Stoker as a fellow student has been important to me for being my outlet while brainstorming. Last but not least I would like to thank all participants that have filled in my surveys. Even though some of them might not have finished the long surveys, all information has been used and was very valuable in order to analyze the data that these participant have given me.

Abstract

A lot of data is abstracted from online surveys, while data validity is related to survey design. The decisions that researchers make within their survey design will influence data outcomes and therefore the decisions that will eventually be made. Privacy and motivation are key elements for measurement errors and by enhancing these elements data validity increases.
This thesis aims to enhance privacy and motivation through better survey design. Better survey design could be accomplished by the provision of a taskbar, no-answer-option and questions concerning sensitive topics. These factors are examined to enhance data validity.

Summary

The main challenge of this study is answering main research question: "How could privacy and motivation be enhanced through better survey design?". In order to give an answer to this question a distinction has to be made what kind of survey designs are used to create higher data validity. This research revolves around adding the taskbar, no-answer-option and exploring more about the influence sensitive questions have on topics like motivation and the respondent's sense of privacy. Before looking into the survey design factors different kind of systematic response tendencies are distinguished to find out more about the background and the existing theories around the topic of common-method variance; a systematic response tendency that is attributable to the measurement method. The measurement methods are the survey design factors and desk research is done on how motivation is influenced by the provision of the taskbar and the no-answer-option. The focus is on ex-ante remedies as they give a more detailed overview about what causes the design factors to influence the outcomes, specialized in preventing rather than curing this problem. The tests gave some contradictory results on both the relationship between motivation and the taskbar as the no-answer-option. More research is done on how the provision of the taskbar influences completion rates and how this changes over the progress in the survey of the respondents. Another element of what role privacy plays if the respondent does or does not have the opportunity not to answer. The data shows that the number of social desirable answer does not increase significantly whenever the no-answer-option is included. The option not to answer is not used significantly more on sensitive questions than the non-sensitive questions, indicating that motivational concerns also influence the number of times respondents used the no-answer-option.
In other words this study tests existing theories and adds value to the discussion on the use of the no-answer-option, taskbar and including sensitive questions.

Chapter 1: Introduction

A lot of marketing decisions that companies have to make, depend on actual data. In order to come to decisions, firms produce data. In terms of marketing decisions a lot of data is created by consumer surveys to find out more about the consumer that is targeted by the company. These consumer surveys are presented to consumers in order to gather information about their needs and opinions. Just like real conversations, the way that questions are phrased could have influence on the answer. In the case of consumer surveys the design of the survey could influence the outcomes; in other words it could lead to misleading and inaccurate conclusions (Campbell & Friske, 1959).

This phenomenon in marketing and psychology is called 'common method variance' (Podsakoff, Mackenzie, Lee & Podsakoff, 2003) and refers to various biases that may affect the validity of data. In this thesis an attempt is made to classify the way in which these false correlations lead to different outcomes of the surveys. The way that a consumer survey is designed is related to different design factors and these design factors could lead to biased results through false correlation. According to the study of Cote and Buckley (1987), the variance in a typical research measure that is caused by systematic sources of method biases is around 25% and differs per subject of interest (Williams, Cote & Buckley (1989). These variances are influenced by all kinds of design factors, this study focuses on three types of designs. The first design factor is adding a taskbar to a basic consumer behavioral survey and the second design factor is supplementing a no-answer-option to a standard response scale. The third design factor is adding sensitive questions concerning delicate topics like drugs, sex and family.

These design factors have influence on privacy and motivation. Sensitive questions lead to misreporting and the extent of misreporting depends on design features of the survey. (Touranqeau & Yan, 2007). This thesis investigates the role of the no-answer-option in enhancing privacy concerns of the respondents. Supplementing a taskbar gives the respondent feedback about their position in the questionnaire(Yan, Conrad, Couper & Tourangeau 2011). Crawford, Couper & Lamias (2001) show that a online survey in which a taskbar is used led to more positive feelings about the participation in the survey than without a taskbar. This thesis does research about enhancing motivation in order to get more accurate reports by the respondents. The main objective of my study is to enhance privacy and motivation through better survey design.

1.2 Problem statement and research questions

The main problem is that certain response tendencies may have a negative influence on the validity of the data outcomes and thus yield potentially misleading conclusions. The validity of the research data depends on the way the consumer survey is designed. The thesis' objective is to identify the ways in which better design influences outcomes. As explained in the introduction of the thesis, there are many different design factors. I focus on the influence of an added taskbar, the no-answer-option and the sensitive questions on privacy concerns and matters of attention/motivation. Therefore the research question of this study is:

"How could privacy and motivation be enhanced through better survey design?"

The first step is to distinguish different tendencies and then to focus on the causes, constructs most prone, as well as the better survey design and statistical methods to deal with it. Giving answer to the question how to enhance privacy and motivation in survey through better design and the effect design factors have on the validity of research data. It refers to more profound marketing problems like the overall validity of marketing research, but this study specializes on the use of the taskbar, no-answer-option and sensitive questions to increase data validity.

1. *What systematic response tendencies can be distinguished?*
2. *How should the effect of the no-answer-option, taskbar and sensitive questions on survey validity be assessed?*
3. *How is attention of importance with the taskbar, no-answer-option and sensitive questions?*
4. *How does a taskbar influence survey response?*
5. *In what way does privacy concern the possibility not to answer?*
6. *How do no-answer-options influence survey response for sensitive questions?*

1.3 Academic relevance

This thesis is about finding ways to enhance privacy and motivational concerns using different kinds of surveys that include methods that might have a result on the data that is produced by the participants of the questionnaire. Thereby the study is trying to verify or falsify existing theories about the design factors at hand: taskbar, no-answer-option and sensitive questions. Besides testing these theories this study makes an effort to add value to the current discussions on these subjects.

Adding value to the common agreement along consumer researchers in both marketing and psychology that common method variance is an acknowledged problem for research results (Podsakoff, Mackenzie, Lee & Podsakoff, 2003) is one of the key elements of this thesis. Contributing is done by quantifying and discussing the role of the survey design at hand. The inducement of this study lies in the findings of Campbell & Friske (1959) finding survey design as a cause for and cure for misleading and inaccurate conclusions. Another assumption that is underlined by this study is the research done by Kahneman and Tversky (1972) who refer to biases in consumer surveys as being a 'cognitive bias', in other words this means irrational judgments that could be triggered by changing design factors.

The work of Yan, Conrad, Couper and Tourangeau (2011) about length of the interview and promised task duration is used as a theoretical foundation on which research is done. Thereby this thesis contributes to the existing literature about common method variance, cognitive biases and other design factors and adds value to study of these specific design factor. This thesis can be seen as a social study to research theories of social scientist like Petty & Cacioppo (1986) that try to explain why motivation and bias influence final results. The work of Tourangeau & Yan (2007) is used as a benchmark to find out how privacy concerns influence results and research is done how the taskbar and no-answer-option enhance the provided concerns on privacy.

Notable to the existing theories are the conflicting results that have been provided by the explorative research on various subjects that are studied in this thesis. For Myers (1985) shows that online survey that provided a taskbar led to more positive feelings about the participation in the survey. Couper, Traugott and Lamias (2001) on the other hand claimed giving feedback about the progress in the survey would display no difference in completion rates . Therefore this study focuses on examining these works and add value to the previous work on the field of creating better survey design to enhance motivation and privacy.

1.4 Managerial relevance

The inducement of this study comes from my political background as well as my interest as a student in finding ways to gather reliable data to form an unbiased opinion. In politics a lot of decisions are made using polls to find public opinions, while survey design has an unambiguous influence on the results (Blumenthal, 2008). Jacobs (2013) also found a lot of methodological flaws in the reformation of the education in Belgium. These kinds of measurement errors are to be acknowledged and this study tries to find out how decision makers can identify these measurement errors. Also this study finds ways for survey designers to cope with the different types of design factors.

From a professional point of view I have seen a lot of marketing decisions at online retailer Coolblue are made based on tracking the consumers' clicks online. Changing the site is a way to change the design of this research and the same happens when a survey is changed; the results change as does the interpretation. For example Jinks, Lawson & Daniels (2003) did a survey on health needs of hospital staff based on surveys on a total population of 2300 hospital employees (response rate= 44%) implicating that initiatives focusing on employees' weight control and taking more exercise are efficient tools to improve health of the hospital staff.

The provision of the taskbar and the no-answer-option to the survey might result in completely different results. Managers should acknowledge the impact of these survey design and make them aware that different methods might lead to different results.

Therefore managers of all kinds of entities, both in the public and private sector should interpret consumer data by the way it is collected. The survey designs that are used in this study (taskbar, no-answer-option, sensitive questions) are applicable for survey designers in order to find the right modus to enhance privacy and motivational concerns. Jääskeläinen et al (2012) developed a checklist which can be used by managers at performance management systems. This study will not develop a checklist, but will contribute the awareness of how design factors have influence on privacy and motivational concerns.

In other words this paper adds to the acknowledgement of the use of design factors and the influence it might have and provides managers tools to find out how certain design factors influence data validity. The purpose is for managers to empathize that choices made on design could influence final results and to underline in what way researchers could cope with issues like privacy and attention.

1.5 Outline of the thesis

The structure is mostly determined by the various research questions. Identification of systematic response tendencies is done through an intensive literature review. This desk research is the literature foundation this thesis needs to make assumptions and answer other research questions. The literature review is completed by exploring how attention and privacy play an important role for the respondents and what possible results a couple of design factors might have on the results and data validity. Also the research questions concerning assessing how the design factor could affect data validity and what types of systematic response tendencies are relevant, are answered. Explaining why a taskbar, no-answer-option and sensitive questions could affect attention is also done by further desk research resulting in hypotheses is described in chapter 2. Subsequently to the theoretical framework the conceptual model is presented and illustrated in chapter 3. Before data analysis, chapter 4 gives space to the methods used to do research are explained.
Besides standard information about how the surveys are made and which questions are used, this chapter is meant to explain how, why and which measurements are used by analyzing the hypotheses in chapter 5. Before starting analyzing is crucial to show descriptives, because between rejected and keeping the null hypothesis all kind of background information could be very valuable. Chapter 5 is all about testing and analyzing the hypotheses that are drawn up in the literature review. After that the conclusions and the recommendations for further research are presented in chapter 6.

Chapter 2: Literature review

In order to find out how design factors like the taskbar and the no-answer-option influence consumer surveys, a qualitative study has to be done in terms of the existing literature that concerns this topic. First the systematic response tendencies are distinguished in greater detail. Secondly the way how the data validity of design factors should be assessed is discussed. After determining how research should be done, the design factors that are studied will be further discussed and the specific effects on the consumers decision making process. A great deal of attention is focused on the relationships between the design factors and the attention and the privacy of the consumer.
Adding a taskbar to the consumer survey will only have effect on the attention (or motivation) of the consumer, but adding the no-answer-option to a response scale has a more complex effect on the respondent. This study makes a distinction between the influences that this design factor has on the respondent as the effects are two-folded. The no-answer-option has both effect the attention, just like the taskbar has, but also it relates to the privacy of the respondent. The respondents might either have a lack of interest in the subject or may not like to state anything about this question due to privacy concerns. Thus choosing 'no answer' in a questionnaire means either the respondent does not care that deeply (attention) or the respondent prefers not to state an answer (privacy). Both privacy and attention are discussed in the literature review.

2.1 Systematic response tendencies

Data can be obtained in different ways, one of them is by surveys. These surveys give the researchers data about the consumer by using applied statistics. The data could contain variance that will be negatively correlated to the validity of the data. Whenever this variance is attributable to the measurement method rather than to the constructs the measures represent, this is called common-method variance (Podsakoff, MacKenzie,Lee & Podsakoff, 2003). Common-method variance is a bias that is caused by the methodology (Richardson, Simmering and Sturman, 2009).

Measurement errors come in two types random and systematic measurement errors. Although both types are problematic, systematic measurement errors are alternative explanations for the observed relationships between measures of different constructs. The main sources of systematic measurement error is method variance that may arise from all sort of response biases as social

desirability, leniency effects, yea- or nay-saying, the halo effect and other effects (Bagozzi & Yi, 1991). Cote & Buckley (1987) and William, Cote & Buckley (1989) found that the amount of variance attributable to method biases differs considerably per subject of the survey. For example, method variance was lowest in the field of marketing (15.8%) and highest at education (30.5%). They also found out that typical job performance measures contained an average of 22.5% method variance, whereas attitude measures contain an average of 40.7%.

The way to distinguish the types of common method variance is by finding out what causes the measurement error. Some method effects are created by the belief of the respondent to be consistent in the responses given, they are called 'common rater effects'. This is a form of social desirability that will be discussed further along this study, when research is done on whether adding a no-anwer-option to the response scale will change depending the sensitivity of the questions. Crowne and Marlowe (1964) did research on social desirability that 'refers to the need of social approval acceptance and the belief that it can be attained by means of culturally acceptable and appropriate behaviours". It is the tendency for the respondent to mask their true opinions in order to show that they are not judged by their opinions.

Another factor of measurement error is caused by item characteristics effects, like when a respondent's interpretation of a question is different because of specific characteristics that this questions possesses. For instance if a survey or questionnaire contains all the same scale formats, this could lead to measurement errors. Tourangeau, Rips and Rasinski (2000) dedicated a study to the scale formats and also studied 'scale anchors', that is the measurement error that is provoked when the survey contains the same anchors (e.g. 'always vs. never' or 'strongly disagree vs. strongly agree'). Another case of item characteristic effect is the 'positive and negative item wording'-effect (Idaszak & Drasgow, 1987) this includes the fact that if a question for example contains positive words it would correlate more strongly towards a positive association to the variable at hand, leading to biased results.

Besides the characteristic effects on items, context effects are also of significant importance to be recognized. It is about the 'interpretation that a respondent might ascribe to an item solely because of its relation to the other items making up an instrument' (Wainer & Kiely, 1987). A well-known effect is the priming effect, which is discussed in the study of Salancik and Pfeffer (1977). It is the saliency of the respondent at a question if the former question was related to the same subject. Also the context-induced mood is discussed plenty in literature, it describes the artificial covariance

that is created for instance by words that respondents relate to gender or ethnic stereotypes, why respondent's mood is affected in a negative way (Peterson, 2000).
The final factor that relates to measurement error is the way that surveys are measured in the same context. In appendix 1 an overview of the most important sources for common method variance. Common method variance is influenced by design factors. A lot of different design factors have already tested on different variables, like response rates on the length of the survey (Yu & Cooper, 1983). Strong evidence have been found that preliminary notification was very effective to increase the response rate (Linsky, 1975). Also monetary incentives in the form of sponsorship were found to have a strong empirical evidence for effectiveness increasing response rates (Kanuk and Berenson, 1975). A review from Harvey (1987) found follow-ups and stamped reply envelopes were also boosting response rates.

2.2 Assessing data validity on design factors

Predicting validity and the maximizing reliability of consumer surveys is a primary goal of this study together with the acknowledgement for managers of the influence of design factors. A lot of attention is drawn to potential remedies to the problem as it is stated in the problem statement. There is one very important distinction to be made in terms of studying solutions to influence of the design factors at hand: ex- ante and ex-post remedies.

Ex-ante remedies for design factors is preventing that the consumer surveys are designed a certain way (e.g. with the taskbar and no-answer-option), so the researchers make a choice whether or not the taskbar and the no-answer-option are of added value to the survey and acknowledge the effect when the surveys are designed without these extra features. Chang, van Witteloostuijn and Eden (2010) discussed remedies to some of the design factors in their article about common method variance. No prior research about the taskbar and the no-answer-option has been found, so a great challenge for this study.

Ex-post remedies is about finding solutions after the research has been done to correct the false correlations due to the design factors in the survey. There are three techniques for testing common-method bias: the confirmatory factor analysis marker technique(CFA), the unmeasured latent method construct (ULMC) and the correlational marker technique (Richardson, Simmering &

Sturman, 2009). Williams, Hartman and Cavazotte (2010) show how this technique is used so that the variance is compensated by this statistical technique. This study focuses on the ex-ante remedies as they give a more detailed overview about what causes the design factors to influence the outcomes, specialized in preventing rather than curing this problem. Preventing the problem is a (financially) much better option than curing the misleading results.

Ex-ante remedies should be about the context measurement, the way that the data results are gathered. In face-to-face interviews the respondents tend to show more social desirability than when they are filling in a survey on the internet (Martin & Nagao, 1989). Leeuw and van der Zouwen's (1988) found that socially desirability bias is even worse in telephone interviews than in face-to-face interviews, so these are not the optimal way to find truthful results. Tourangeau and Smith (1996) found that self-administration eliminated the gap between the reports of men and women, decreasing the average number of sexual partners reported by men and increasing the average number reported by women.

Also the way surveys are set up in terms of item characteristics is an important focus of this study. On the one hand scale format should not be the same, as well as the scale anchor should not be the same (Tourangeau, Rips & Rasinski, 2000). On the other hand in order to prevent contextual effects the questions are presented by the Likert-scale and done in the same anchor as the results will not be different for every single ordinal question. Bradlow and Zaslavsky (1997) did a similar study to the no answer option and they recognized that this kind of approach has some typical shortcomings, like there is usually not a suitable treatment for item non-response.

In other words this study focuses on research about the validity of the data when using the taskbar and no-answer-option. First of all this study uses a quite long and extensive consumer survey as platform for supplementing or leaving out the taskbar and no-answer-option. The length of the consumer survey has significant effect on attention level and on the added value of the taskbar. In general the assumption is that the longer the survey, the higher the non-response rate and the lower amounts of breakoffs (Conrad, Couper, Touranqeau & Petytchev, 2010). In order to reduce potential non-response bias, survey practitioners should try to keep their surveys as short as possible, but to research the added value of the taskbar a long survey is needed.

2.3 Motivation

A great deal of this study revolves around the human mind and to find out how the consumers' mind

reacts to design factors. The elaboration likelihood model (ELM) shows how the mind is part of the 'elaboration continuum'. A distinction is made between low elaboration and high elaboration. The ELM distinguishes two important routes: the peripheral and central route.
The central route is the case with respondents that are motivated and use high thoughts to choose their answers. The questions the respondent faces are considered by using logic and are answered in a sensible matter. The peripheral route is a route when the respondent shows low motivation and therefore uses preexisting ideas, therefore the respondent is more reliant on biases (Petty & Cacioppo, 1986). This would have considerable effect on the respondents answer and therefore also on the variance and final outcomes of the surveys.

The elaboration likelihood model is not the only model that describes and explains how people receive and process information. Another model is the Heuristic-Systematic Model of Information Processing (HSM) , another widely excepted communication model. The HSM also distinguishes two main streams of processing by respondents, namely heuristically and systematically (Chaiken, 1980). The HSM is closely related to the ELM. HSM specifically examines "validity seeking" persuasion settings concerning people's motivations within the social environment The limitation of HSM exists in the inability to define the specific motivations of persuasion, which is why the HSM is expanded to illustrate that heuristic and systematic processing can "serve defense-motivation, the desire to form or defend particular attitudinal positions, and impression- motivation, the desire to form or hold socially acceptable attitudinal positions" (Eagly & Chaiken, 1993). The ELM is therefore more suited to be used in order to obtain the common-method variance and false correlations of which the research data may suffer.

Another ELM-related theory can be found in the human information processing theory, which makes a distinction between deep processing and shallow processing by Craik & Lockhart (1972). This theory describes memory recall. For instance depth of processing is related to the memory traces, where as deep processing results in a more durable memory trace.

The length of the survey relates to the elaboration of the respondent. Research is done by using a survey that takes approximately 25 minutes to complete, it is thus a relatively long consumer survey to test the attention of the respondents over a significant number of minutes. Testing the taskbar has already been done, but resulted in a lot of different findings. The assumption among survey methodologists is that respondents want to know their position in a survey, so giving progress feedback seems to have a positive impact on completion rates (Conrad, Couper, Tourangeau & Petytchev, 2010). Myers (1985) shows that online survey that provided a taskbar to give the

respondent feedback about their position in the questionnaire, led to more positive feelings about the participation in the survey than without a taskbar. But there were also studies that claimed giving feedback about the progress in the survey would display no difference in completion rates (Couper, Traugott & Lamias, 2001). Boltz (1993) assumes that people allocate mental resources for a task based on how long they expect the task to be. In other word if a respondents' expectations change when the feedback indicates the task will take longer than expected, completion rate will decrease, because respondents are feeling disappointed by their feedback. Kahneman, Knetsch & Thaler (1991) explain how losses are often more painful than comparable gains. This principle is also in line with the research done by Boltz (1993) suggesting that people allocate mental resources for a task based on how long they expect the task to last. If the duration of completing the survey is consistent with the duration that was indicated in the invitation Yan, Conrad, Tourangeau & Couper (2007) it reduces the number of breakoffs, therefore this study have communicated a fair amount of time that was expected to finish the questionnaire (appendix 3).

Also the surveys have a couple of similar questions that are both positively and negatively asked. To enhance motivation at long surveys Kaplan and Saccuzzo (2009) found out that by alternating positively and negatively worded questions keeps respondents 'on their toes', plus it can identify those respondents that are not processing the questions. In order to learn more about attention, this study wants to find out more about the diversity of answers given by the respondent. Diversity of answers shows how attentive the respondent is about answering questions. This study expects to find a reduced elaboration over time spend to fill in a long survey. Therefore diversity of answers are expected to reduce every question screen the respondents complete in their survey.

As cognition is an individual trait, respondents choose to take the central or peripheral route of processing the survey. This study does research on design factors that influence attention in the form of completion rates and elaboration. Completion rates show the amount of breakoffs/ non-response and indicate whether or not feedback in the form of a taskbar is effective. That is the reason why this study also does research on all respondents that did not finish the survey as well.

Hypothesis:

1a. Diversity of answers is significantly higher with the provision of the taskbar.

b. Diversity of answers is significantly lower in the second half of the survey compared to the first half of the survey.

c. Diversity of answers is significantly higher in the second half of the survey with the provision of the taskbar.

2a. The provision of the taskbar has a significant positive relation to the completion rates.

b. The number of breakoffs significantly decrease in the second half of the survey with the provision of the taskbar.

Motivational factors include the personal relevance of the message topic, accountability, and a person's "need for cognition" (their innate desire to enjoy thinking). The personal relevance of the message topic and the accountability are difficult to measure, although they are stimulated by choosing a message topic, during the experiment that is relevant. Though the personal relevance cannot be verified in numbers (easily) and could be different for every individual respondent. The same restrictions are relevant to the subject of accountability. A person's need for cognition could be effected by a taskbar or no-answer-option. The need for cognition (NFC) is discussed by Petty and Cacioppio (1982) and concerns the extent of engagement and enjoyment of effortful cognitive activities. Another definition is that an individual's NFC is "a need to understand and make reasonable the experiential world" and "a need to structure relevant situations in meaningful, integrated ways" made by Cohen, Stotland & Wolfe (1955). Dole and Sinatra (1988) use the elaborated likelihood model to make a clear distinction between high and low elaboration. Need for cognition is an individual trait that is the respondent that evaluates ideas, analyzes problems and use effort to come to their conclusions according the central route. The peripheral route of low elaboration contains the respondent that shows a low need of recognition and have low motivation to put thought and effort in their analysis to answer the survey questions. These respondent are more sensible to answer the research question using bias and irrationally.

Effort-based question increase the probability of defaulting; choosing not to choose. These effects matter most when the stakes are small or when people do not fully compensate the effort of answering the survey question with the importance of making the decision. Fixed questions have a tendency to increase completions. They observed that much of the abandonment occurred on questions requiring users to type free text (McKenzie, Liersch & Finkelstein, 2006). A respondent could have a higher default rate due to effect of rational inaction, but given the large effects of defaults in consequential domains of sensitive questions, it seems like giving people the option of not answering might be less likely due to an optimal allocation of effort instead of protection private matters.

John, Naumann & Soto (2008) did research on the so-called 'Big-Five Traits' that is introduced as the five-factor-model of personality traits by Costa & McCrea (1992) and found a way to find solutions to some conceptual issues. This study inspired me to use conscientiousness as an indicator of how serious the respondent take questionnaires and thereby the motivation.

Hypothesis:

3a. Conscientiousness is positively related to choosing not to answer with sensitive questions.

b. Conscientiousness is positively related to choosing not to answer without sensitive questions.

2.4 Privacy

Privacy issues result in respondents being reluctant to give a truthful answer, this leads to misreporting. Misreporting could be done in two separate ways, namely by under- and overreporting. Wish, Hoffman and Nemes (1997) studied self-reports on drug abuse and found out that between 30-70% of the cocaine users were lying about ever using the drug, in other words they were underreporting and this leads to defective survey results.
Overreporting means that due to privacy concerns higher measures are observed by respondents answering questions about desirable acts. Overreporting happens a lot less in terms of sensitive questions, but to give an example respondents are more likely to report the use of their seat belts more often than they actually do (Stulginskas, Verreault & Pless, 1985) and respondent tend to report higher efforts on the conservation of energy (Fujii, Hennessy, & Mak, 1985).
Underreporting means exactly the opposite, showing lower measures when the respondent is asked about doing undesirable acts like drugs and alcohol abuse.
According to Joinson, Woodley and Reips (2007) respondents are more likely to answer sensitive questions with 'prefer not to state' if the respondents have the feeling that their anonymity is compromised, for example by questions concerning highly person matters. The use of the 'I prefer not to state' as answer to a sensitive question is methodologically similar to the use of the no-answer-option (Knapp & Kirk, 2003). Also the intensity of the questions is important to the measurement of the influence on validity of the 'no-answer'-possibility (Dinner, Johnson, Goldstein & Liu, 2011). There are all kinds of personal/ sensitive questions that can be added to a survey; questions about drugs, voting, salary, family relations, sex, etcetera.

Hypothesis:

4. Sensitive questions cause respondents to choose significantly more no-answer-options than non-sensitive questions.

Adding a no-answer-option to the response scale ensures that the respondents have the option to keep up opinions or attitudes private. this is a form of social desirability, where people do not want to be embarrassed by their answers. The methodological findings suggest that socially desirable responding in surveys is largely contextual depending both on the facts of the respondent's situation and on features of the data collection situation such as the degree of privacy it offers (Tourangeau & Yan, 2007). Johnson and Goldstein (2003) have concluded that choosing not to answer might occur caused by these three reasons: reference dependence, whereas this answer determines that strength or weakness of the answers on the following questions. Another reason is the attention level that the respondent wants to provide into answering this sensitive question and the final reason is because of social desirability.

A form of social desirability was optioned by Mckenzie, Liersch & Finkelstein (2006) which is called 'implied endorsement', explaining that construction of the question makes the respondent think what the most desired answer should be. Implied endorsement was expected to have a large influence when the survey questions contain social options of significant importance to the respondent (Dinner, Johnson, Goldstein & Liu, 2011).

Tourangeau & Yan (2007) concluded there is evidence that asking sensitive questions boosts item non-response and reporting errors. A strong relationship between the sensitive questions and misreporting was found. The article finds socially desirable responding in surveys is largely contextual, depending both on the facts of the respondent's situation and on features of the data collection situation, like design factors.

Choosing specifically for surveys presented by internet is the most efficient way of achieving higher levels of self-disclosure and an increased willingness to answer sensitive question (Tourangeau, 2004). When data is collected by computer-aided interviews respondents show more signs of self-disclosure, for instance about health related problems (Epstein, Barker & Kroutil, 2001). De Leeuw (1992) found out privacy concern had a significant relationship to socially desirable answers as face-

to-face-interviews resulted in an unusual large amount of socially desirable answers. Face-to-face interviews or by telephone might also results in more quality answers as the questions could be explained a bit more and the interviewer could motivate the respondent to a higher extent according to Dillman, Smyth and Christian (2009).

Hypothesis:

5a. Respondents report significantly more social desirable answers on sensitive questions without the no-answer-option.

5b. Illegal sensitive questions cause for significantly more number of times respondents chose not to answer than legal sensitive questions.

Chapter 3: Conceptual model

The conceptual design display two important elements: (a) a collection of concepts that is crucial to the study and (b) show the relationships between the concepts. This study examines two major concerns; motivation and privacy (dependent variables). Both concerns are investigated through using three different survey design methods; taskbar, no-answer-option and sensitive questions (independent variables). In order to create clear overview of all relationships two different conceptual models are used, because this study is possessed by two dependent variables.

A direct effect between independent variables (provision of taskbar and sensitive questions) and the dependent variable (motivation) is tested by hypothesis 1 concerning the influence between the diversity of answer, an indicator of motivation, and supplementing the taskbar as a design factor. The second hypothesis tests other indicators of motivation; completion rates and breakoffs (graph 1).

1. **Concerning the task bar**

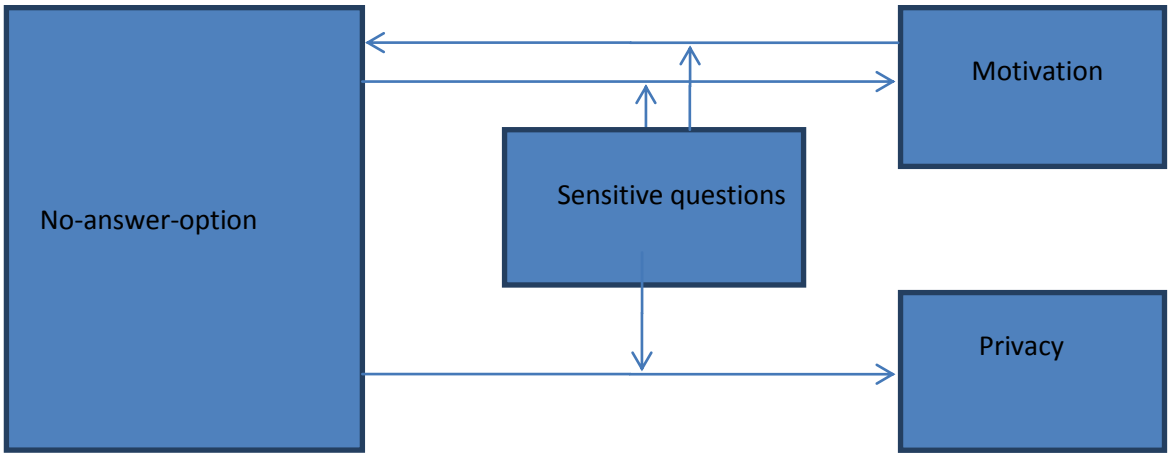
Graph 1

The no-answer-option is the independent variable that is tested to have an effect on motivation in hypotheses in the third hypotheses, whereas the relationship between conscientiousness as indicator for motivation and the provision of the no-answer-option. Hypothesis 5 insinuates a relationship between socially desirability (an indicator of privacy) and the provision of the option not to answer, while the fourth hypothesis tries to indicate that adding sensitive questions have an influence on the use of the no-answer-option. This leads to a direct effect between the design factors and therefore the sensitive questions have an interaction effect on the aforementioned relationships between the no-answer-option and the dependent variables. Instead of an indirect effect the literature review assumes the effect differs between the dependent variables. Sensitive questions is a control mediator on the design factor to enhance privacy and motivational concerns.
The reason why this conceptual model also includes a reverse effect between motivation and the option not to answer, is because of hypothesis 3, finding out more about the relationship between conscientiousness (indicating motivation) and the no-answer-option. In this hypothesis the no-

answer-option is the dependent variable in order to make sure whether or not motivated respondents react differently to the supplement of this option.

Hypothesis 3a and 3b gives an indication of the strength of the relationship and the interaction effect, therefore the sensitive question is also a mediator to the influence between motivation and the no-answer-option (graph 2).

2. Concerning the no-answer-option/ sensitive questions



Graph 2

Chapter 4: Methodology

In order to analyze the hypotheses it is key to understand the data and the underlying methods used to abstract whether the null hypothesis is rejected or retained. This chapter reports all important aspects of methodology and data description. Survey design is the foundation for all tests, while randomization and the invitation tell something about how the samples are approached. The invitation for instance influences privacy concerns and random samples is a general assumption for most tests.

The construction of the measures are crucial to explain the underlying aspects of each analysis. Hypotheses contain several concepts and in order to make these concepts operational measures are determined and defined. Besides the measures it is important to the reliability of the results to find out more about details like outliers, means and standard frequencies. This kind of descriptive data adds annotations that might be important for the conclusions drawn after analyzing the tests in chapter 4.

4.1 Survey design

The survey contains questions around the topic of consumer behavior. Behavior in general and in the workplace, spending habits, brand-related or about 'hot topics' like the environment, health and so on. In other words the standard survey contains all non-sensitive questions that will not evoke many socially desirable responses.

In order to do research on the provision of the no-answer-option (NA-option), four surveys are needed as the provision of sensitive questions (about drugs, sex, temptation, etc.), the taskbar and the no-answer-option are used to seek for optimal results (table 1).

Survey 1	Survey 2	Survey 3	Survey 4
- no taskbar - sensitive questions - NA-option	- taskbar - non-sensitive questions - NA-option	- taskbar - sensitive questions - no NA-option	- no taskbar - non-sensitive - no NA-option

Table 1

Survey 2 and 3 are extended by 21 sensitive questions. These questions are used as a tool to convey in what way the option 'not to answer' is used by the respondent. Privacy is a very important

component (hypothesis 5a, b) and by adding the sensitive questions, privacy boundaries are examined.

Sensitive questions verge on certain boundaries, but the boundaries that hypothesis 5b focuses on are linked to whether it is legal or not. The division of these sensitive questions on legality of the possible answers (table 2).

Sensitive questions:

Legal	Illegal
• I would buy a scarf made of animal fur • If I could invent something to secretly cheat in my relationship, I would do so • I would like to live in Holland without the people that were born in another country • I wish to have sex with people I feel attracted to, even if I cheat in a relationship • I lie about the amount of people I have had sex with • I would lie about somebody else to make myself look better • I wish I had different parents • Sometimes I think about my own suicide • I lie about my salary to friends • I wish I did not have family	• I look forward to parties where I can do (illegal) drugs • I wish to break the law if nobody would find out • I would like to disobey the law if I could talk my way out of a ticket • If I could steal something knowing I won't get caught, I would do so • I have used hard drugs • Neglecting to put a seatbelt on if I know I won't get caught, seems like something I would do • If I had the possibility to drive faster than allowed without getting caught, I would definitely do so • I would drive of have driven a car with too much alcohol allowed in my blood • I have peed in public (in the wild or somewhere without using the toilet) • I have been rude to a cop and called him/her names • I have stolen things from people close to me

Table 2

The surveys in this study use fixed questions with a ordinal Likert-scale. In chapter 2 the literature about the content of the questions have led this study to contain questions with a lot of different topics to get the respondents attention. Within these topics a couple of questions that have a slightly altered questions to keep the participants acute.

4.2 Randomization

In statistics, random samples are of great importance to make sure realistic data is used. Therefore the sheet of appendix 2 is used to use a large network of very different people for the study's benefit. The 'Oxford English Dictionary' defines randomness as "Having no definite aim or purpose; not sent or guided in a particular direction; made, done, occurring, etc., without method or conscious choice; haphazard." This concept of randomness is used in appendix 2 as every individual inside a diverse network (students, colleagues, friends, family, members of the gym, tennis, football and a lot of different political parties) has been given the same message asking for their participation to the survey/questionnaire.
All names are ordered alphabetically and the whole group is divided by four groups; so all four group are given different links to the four different surveys to improve randomness.

4.3 Invitation

The message they received as invitation to the survey was based on a realistic forecast of time that most of the respondents would need to fill in the complete survey, approximately 20-25 minutes. Indicating the time the respondents need to finish the survey is extremely important. The amount of people that start the survey would have been greater if the invitation stated it would take 10-15 minutes, but the completion rate would probably be much lower, because if respondents experiencing discouraging information after 15 minutes (by the provision of the taskbar) or feel like they are misled by the invitation that is taken to heart with a greater degree than encouraging information. Literature review explained how losses are often more painful than comparable gains, in other words it is better to be fair or a little bit precautious about the amount of time it generally takes for the respondents to complete the survey. Respondents allocate mental resources for a task based on how long they expect the task to last. If the duration of completing the survey is consistent with the duration that was indicated in the invitation it reduces the number of breakoffs.
Another important component of the invitation (appendix 3) is naming the fact that all answers given

in the survey will not be used in public and won't be visible to anyone, not even the researcher himself to ensure respondents of their privacy.

4.4 Measures

This study revolves around the concepts of privacy and attention. The way to make these concepts ready for operational use is crucial to the final outcomes. Privacy is tested by adding sensitive questions to the survey. Sensitive questions lead respondents to make the following decision: whether to answer social responsibly, honestly, quit the survey or use the no-answer option.
The four options are analyzed, but before they are tested these four options have to be measurable. The number of break offs and incomplete survey are easily measured using frequency tables, just as the number of no-answer-option. Defining honesty or social responsibility takes a profound approach. Honesty is connected to social responsibility, as respondents chose to answer sensitive questions in a way that they think these questions should be answered. In other words if we can measure social responsibility we can say something about how privacy concerns have an effect on the sincerity of the answers.

Two out of four surveys included in total 21 sensitive questions. These sensitive questions have a socially desirable answer, actually they have two. For instance if the question states: 'I have used hard drugs', the only two social desirable answers are 'totally disagree' and 'disagree'. All other answers are not social desirable. Because of this we can state for every answer to every sensitive questions a 0 or 1. Using this standard or rule a new SPSS-file can be created with all zeros and ones with the sensitive questions as nominal variables.

Measuring attention calls for a different approach. Attention is the way the respondents processes the questions and takes the energy and time in answering the questions. If attention drops respondents tend to read the questions more quickly and don't put much effort in answering the questions. Reading the question faster and less precise results in misunderstanding and misinterpreting the questions and therefore in strange unexpected answers. The questions are not designed in order to find these mistakes and conclude whether or not the question is misread.
Putting less effort in answering the questions results in giving the same answers to different questions in order to finish the survey faster and therefore not using 'deep processing'. Conscientiousness denotes being thorough, careful, or vigilant and the surveys includes three questions indicating this trait

and if a respondent has the desire to do a task well. The results of the questions are used to find correlation between the parameter 'conscientiousness' and the no-answer-option. The survey contains three questions concerning conscientiousness:

- I see myself as someone who does a thorough job (positive relationship to conscientiousness)
- I see myself as someone who can be somewhat careless (negative relationship to conscientiousness)
- I see myself as someone who tends to be lazy (negative relationship to conscientiousness)

The results of these three questions give an indication of how conscientious a respondent values him- or herself and can be seen as an indicator of motivation.

Social desirability is difficult to measure as it is a matter of values and norms the respondents have. Cultural differences could play a role as do religious influences have effect on the norms and values of the respondent. For instance wearing a scarf made of fur is a lot more shocking to one respondent than another depending on culture or religion. This accounts for all sensitive questions. To find out whether or not respondents are social desirable answers, this study uses to the law, norms and values are considered as mainstream. Therefore all other answers than the social desirable ones ('totally disagree' and 'disagree') are considered as undesirable. Choosing not to answer or a neutral answer at sensitive questions shows either lack of motivation or the respondent is concerned about his privacy. If he is concerned about privacy and he probably has something to hide, he can answer desirably or choose the no-answer-option. Because of this all other answers beside the social desirable ones are coded as social undesirable ('totally agree', 'agree', 'neither disagree nor agree' and 'no answer'). Coding desirable answers with a '1' and undesirable answers as a '0' gives the possibility to test the hypothesis about whether or not the provision of the option not to answer enhances respondents to answer more honestly.

4.5 Descriptive statistics

Research is done over four different surveys. The surveys do not include personal questions about age, sex or residence (due to privacy concerns). Hence, no descriptive statistics are presented for these variables.

The number of surveys including sensitive questions (113) is lower than the number of surveys without sensitive questions (136). The number of surveys including a taskbar (131) is higher than the number of surveys without a taskbar (118). The number of completed surveys is explained in table 3.

180 of the 249 respondents that started the survey, finished the survey, that is a completion rate of 0,723.

Survey 1:

58 respondents, 16 incomplete

Survey 2:

55 respondents, 14 incomplete

Survey 3:

76 respondents, 20 incomplete

Survey 4:

60 respondents, 20 incomplete

		Frequency	Percentage
Complete	No	69	27,7
	Yes	180	72,3
	Total	249	100,0

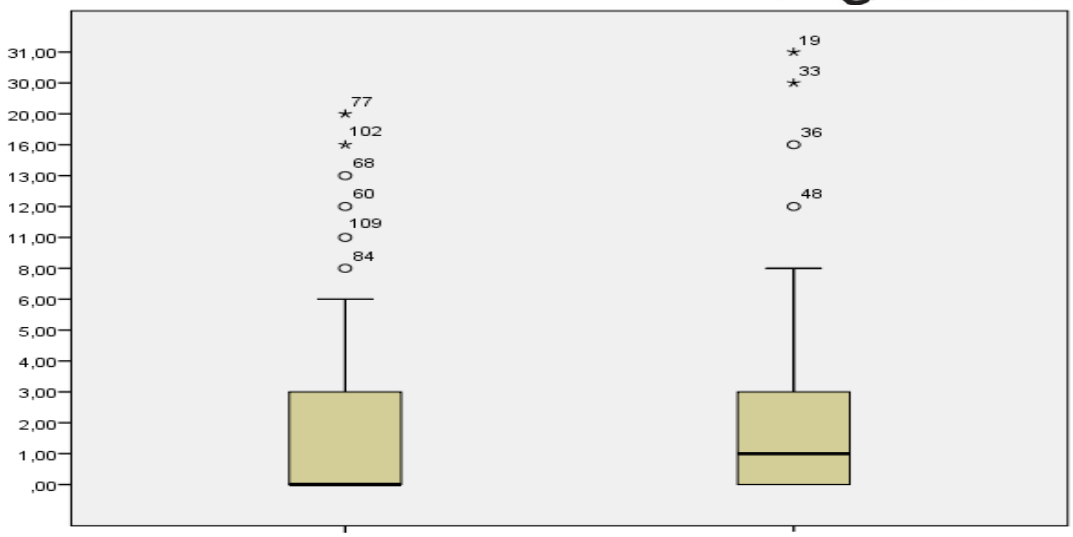
Table 3

The breakoffs occur for the largest part in the first half of the survey. Almost 90% of all breakoffs occur in the first half, while only 3,6% of all respondents end the questionnaire in the second half. The number of times respondents used the no-answer-options differ a lot and the risk of outlier is therefore large, see the respondents that use the option more than 20 times. While it only accounts for 13% of all the respondents that have stopped the survey before finishing, see table 4.

	Frequency	Percentage
Second half	9	13%
First half	60	87%
Total	69	100%

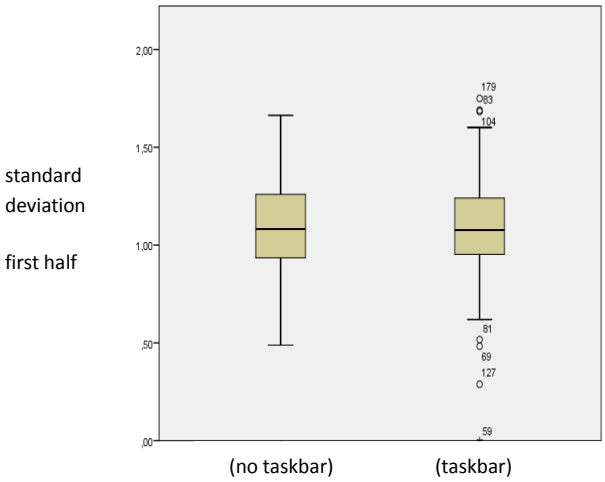
Table 4

The fourth hypothesis states that sensitive questions might increase the number of times respondents chose not to answer. The boxplot (graph 3) shows the difference between the survey with and without the sensitive questions. Remarkable is that the median and maximum of the number of 'no-answers' is obviously higher when surveys contain these sensitive questions.

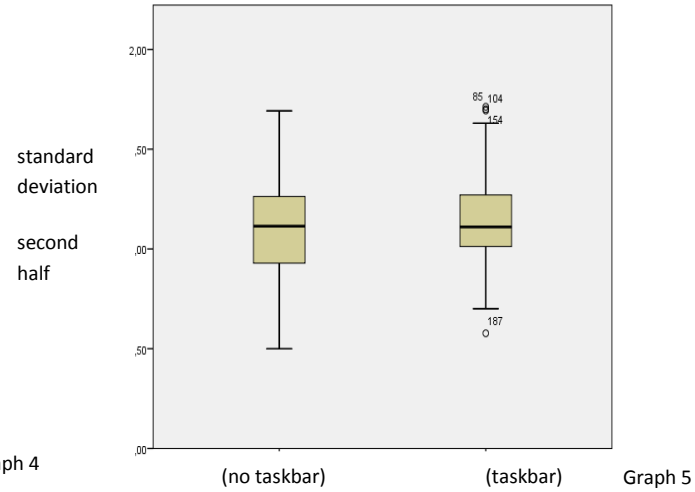


Graph 3

This study also does research on how motivation differs as the respondent progresses. The standard deviation, the variation of answers is compared between the first half (graph 4) and the second half of the survey (graph 5) to find out more about respondent's attention over time.



Graph 4

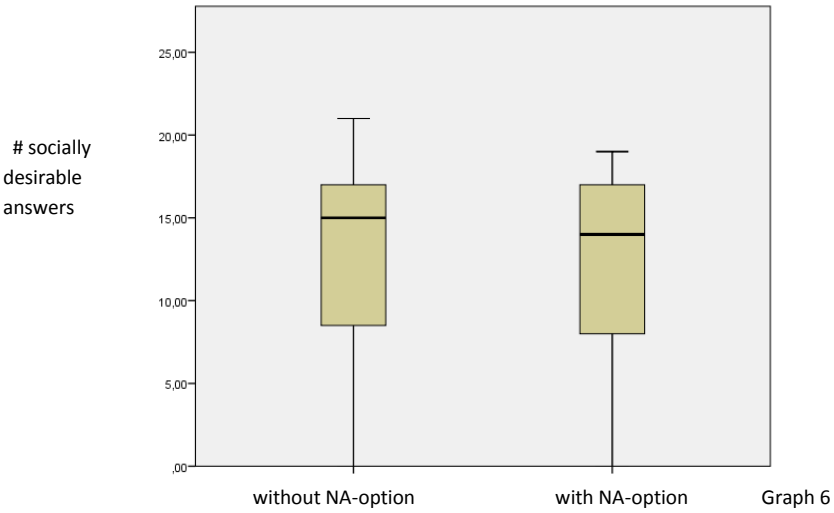


Graph 5

The medians are almost equal at both halves and with or without taskbar. The range between the maximum and minimum is smaller if the taskbar is added. This might indicate a lower amount of motivated respondents when a taskbar is added, although the number of outliers is significantly larger if the taskbar is provided.

The fifth hypotheses states that the no-answer-option could affect the respondent on giving more social desirable answers. The box plot below (graph 6) shows the difference between the number of socially desirable answers given to sensitive questions with and without the no-answer-option are

shown.



Graph 6

Note the smaller range (difference between the maximum and the minimum) when the no-answer-option is provided. Also the median is lower whenever there is an option not to answer. This plot strengthens the expectation that providing a respondent a no-answer-option might enhance privacy, because the respondent is less prone to give social desirable answers.

Chapter 5: Analysis

This chapter is restricted to tests done on all hypotheses that have been discussed in chapter 2 and the measures and methodology defined in chapter 4. Setting up the null and alternative hypothesis is done by a short introduction to the test. This introduction also provides information about the utilization to prove whether the null hypothesis is rejected or retained.

Hypotheses

1a. Diversity of answers is significantly higher with the provision of the taskbar.

In order to test this hypothesis two groups (means) have to be compared. Only means from survey 2 and 4 are used, because sensitive questions give an extra non-ordinal option for the respondent that influences diversity of answers. The standard deviation of each respondent is measured, in order to find out how the respondent's attention reacts on the provision of the taskbar. Therefore two means are compared; one survey including a taskbar and the other without a taskbar.

These observations result in an alternative hypothesis that claims that the diversity is significantly higher when the taskbar is provided (Ha: $\mu 1 > \mu 2$), which is an one-sided test. Comparing means is done by an independent-T-test, results in the following SPPS-data in table 7 and 8.

Diversity of answers

	N (respondents)	Mean	Std. Deviation	Std. Error Mean
No taskbar	60	1,0579	,22688	,02929
Taskbar	55	1,0966	,30097	,04058

Table 6

T-value	Degrees of freedom	Significance (2-tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference	
					Lower	Upper
-,782	113	,436	-,04234	,04913	-,13967	,05498
-,773	100,059	,441	-,04234	,04959	-,14071	,05602

Table 7

T-value of -0,862 is insignificant as 0,391/2 = 0,196 and 0,196 > 0,05.

This outcomes means that there is no significant higher diversity of answers when the taskbar is added to the survey. The expected higher standard deviations are found, but the difference is not significant, so there is no support for this hypothesis.

1b. Diversity of answers is significantly lower in the second half of the survey than the first half

This comparison is made between the diversity of answers in the first and second half by a paired t-test. Diversity of answers is measured by the standard deviation of each respondent. The means between all respondents that have both a standard deviation in the first as in the second half of the survey (n = 189) are compared. Table 8 shows the means of both halves, while the paired difference is measured in the paired t-test in table 9 where the two-sided test results are shown.

	Mean	***Number of respondents***	***Std. Deviation***	***Std. Error Mean***
Standard deviation first half	1,1040	189	,21065	,01532
Standard deviation second half	1,1290	189	,23686	,01723

Table 8

	Mean	***Std. Deviation***	***Std. Error Mean***	***T-value***	***Degrees of freedom***	***Significance (2-tailed)***
Standard deviation first vs. second half	*-,02506*	*,14349*	*,01044*	*-2,401*	*188*	*,017*

Table 9

T-value is -2,401 at a significance level of 0,017 < 0,05, which means that this hypothesis and a significant difference is found between the first and second half. This results shows that attention levels are found to be lower at the second half of the survey of the survey.

c. Diversity of answers is significantly higher in the second half of the survey than the first half with the provision of the taskbar.

This hypothesis is based on the standard deviations of 1b, only we know use the difference (standard deviation of the second half minus the standard deviation of the first half). By testing two unpaired groups with and without the taskbar this hypothesis uses independent-T-test to find out if the attention (diversity of answers). This expectation is that the feedback of the respondent's progress

leads a higher attention in the second half of the survey. Hence, this test is one-sided.

Population proportion	Sample size	Count of successes	sample proportion

SPSS-results of the independent-T-test on the first half (table 10 and 11):

	Number of respondents	Mean	Std. Deviation	Std. Error Mean
no taskbar	80	,0134	,18290	,02045
taskbar	98	,0458	,10844	,01095

Table 10

T-value	Degrees of freedom	Significance (2-tailed)	Mean Difference	Std. Error Difference
-1,468	176	,144	-,03244	,02209

Table 11

T-value is -1,468, but insignificant as 0,144/2 = 0,072 and 0,072 > 0,05. In other words there is no significant higher level of attention/ diversity of answers in the second half of the survey compared to the first half when the taskbar is provided. Literature review discussed the doubts that some academic researchers have about this assumption. This results strengthens these second thoughts about assuming that a progress feedback automatically increases attention or motivation to answer the rest of the questions.

2a. The provision of the taskbar has a significant positive relation to the completion rates.

The provision of the taskbar is measured by using two separate population proportions, which illustrate differences in completion rates. A completion rate is a population proportion, which is p. This unknown parameter is estimated by the statistic called 'sample proportion'. The sample sizes (n) are respectively 131 and 118, while the number of successes are respectively 98 and 82 (tables 12 and 13).

	Breakoffs	completed	Total
No taskbar	36	82	118
Taskbar	33	98	131
Total	69	180	249

Table 12

p1 = taskbar	n1 = 131	X1 = 98	$\hat{p}_1$ --> X1/n1= 0,748
p2 = no taskbar	n2 = 118	X2 = 82	$\hat{p}_2$ --> X2/n2= 0,695

Table 13

The proportion test that is used is one-sided as the taskbar (proportion 1) gives respondents feedback which could lead the respondents to have better knowledge about how to allocate their efforts to might finish more often: p1 > p2.

In order to perform this test the Z-value has to be found:

$Z = (\hat{p}_1 - \hat{p}_2) / \sqrt{\hat{p}(1-\hat{p})}\sqrt{(\frac{1}{n1}+\frac{1}{n2})}$

$\hat{p}$ = (X1 + X2) / (n1 + n2) = 180/249 = 0,723

$Z = (\hat{p}_1 - \hat{p}_2) / \sqrt{\hat{p}(1-\hat{p})}\sqrt{(\frac{1}{n1}+\frac{1}{n2})}$

$Z = (0{,}748) - (0{,}695) / (\sqrt{0.723(1-0.723)}\sqrt{(\frac{1}{131}+\frac{1}{118})}) = 0{,}735$

0,735 at a significance level of 0,05 gives a P of 0,77, which means that p1 is not significantly larger than p2: (1-0,77)/2 > 0,05, which means that provision of the taskbar does not have a significant positive effect on an increase of completion rates.

b. The number of breakoffs significantly decrease in the second half of the survey with the provision of the taskbar.

Comparing the number of breakoffs (table 14 and 15) in the first and second half of the survey leads to a proportion-test.

	2nd half	1st half	Total
No taskbar	5	31	36
Taskbar	4	29	33
Total	9	60	69

Table 14

Population proportion	Sample size	Count of successes	sample proportion
p1 = first half	n1 = 60	X1 = 29	$\hat{p}_1$ --> X1/n1 = 0,483
p2 = second half	n2 = 9	X2 = 4	$\hat{p}_2$ --> X2/n2 = 0,444

Table 15

Expected is that the number of breakoffs significantly decrease in the second half when the taskbar is provided (proportions 1 and 2).

$Z = (\hat{p}_1 - \hat{p}_2) / \sqrt{\hat{p}(1-\hat{p})}\sqrt{(\frac{1}{n1}+\frac{1}{n2})}$

$\hat{p}$ = (X1 + X2) / (n1 + n2) = 33/69 = 0,478

$Z = (\hat{p}_1 - \hat{p}_2) / \sqrt{\hat{p}(1-\hat{p})}\sqrt{(\frac{1}{n1}+\frac{1}{n2})}$

$Z = (0{,}483 - 0{,}444) / (\sqrt{0.478(1-0.478)}\sqrt{(\frac{1}{60}+\frac{1}{9})} = 0{,}437$

0,437 at a significance level of 0,05 gives a P between 0,67 which means that p1 is not significantly larger than p2: (1-0,67)/2 > 0,05, which means that provision of the taskbar does not have a significant positive effect on an increase of breakoffs in the first half of the survey.

Conclusion is that both halves show no significant differences between number of breakoffs. The power of the provision of the taskbar does not lead to siginicant decrease of the number of breakoffs over time a respondent fills in the questionnaire if the first and the second half are compared.

3a. Conscientiousness is positively related to choosing not to answer with sensitive questions.

This hypothesis is tested by the correlation (r) between conscientiousness (x) and the number of no-answers given by the respondents at the sensitive questions (y).

Survey 1 contains sensitive questions and the no-answer-option, so those survey-results are used to find out more about the relationship between motivation and choosing not to answer (table 17). The number of times a respondent chose not to answer was multiplied by 135/156. The number of sensitive questions (21 in total) and non-sensitive questions (135 in total) are compared to only 135 non-sensitive questions in hypothesis 3b. In order to find a realistic comparison the total amount of no-answer-options chosen have to be related to the total amount of questions. The number of

respondents (n) is equal to the amount of completed surveys that also answered the questions concerning conscientiousness (n=41).

		Conscientiousness of the respondents	Number of no-answer-options chosen
Conscientiousness of the respondents	Pearson Correlation	1	,136
	Significance (2-tailed)		,396
	N (# respondents)	41	41
Number of no-answer-options chosen	Pearson Correlation	,136	1
	Significance (2-tailed)	,396	
	N(# respondents)	41	41

Table 17

Correlation shows the direction of the linear relationship between two variables. These quantitative variables are conscientiousness and number of times the respondents choose not to answer. Correlation (r) measures the relationship.

The Pearson Correlation measures a positive relationship, meaning that motivation/ conscientiousness increases the number of times respondents choose not to answer. This relationship is nevertheless very weak and insignificant (significance 0,396/2 = 0,198).

b. Conscientiousness is positively related to choosing not to answer without sensitive questions

This hypothesis tests the survey without the sensitive questions (survey 2). Again the Pearson Correlation is used as method to measure the relationship between variable x and y. Conscientiousness is variable x, while the number of times a respondents choose not to answer is y (table 18). The number of respondents (n) is equal to the amount of completed surveys that also answered the questions concerning conscientiousness (n=40).

		Conscientiousness of the respondents	Number of no-answer-options chosen
Conscientiousness of the respondents	Pearson Correlation	1	-,149
	Significance (2-tailed)		,360
	N (# respondents)	40	40
Number of no-answer-options chosen	Pearson Correlation	-,149	1
	Sig. (2-tailed)	,360	
	N (# respondents)	40	40

Table 18

Notable to the relationship between x and y, it shows a prominent different relationship. The linear relationship is weak (significance 0,360), but measures a negative direction (- 0,149) of the correlation. In other words conscientiousness shows a negative relationship for choosing not to answer, if sensitive questions are excluded of the survey. This relationship is not significant(0,36/2 = 0,18) at a significance level of 5%.

The results of hypotheses 3a and 3b indicate that surveys with sensitive questions leads a conscientious respondent to use the no-answer-option more often, compared to surveys without sensitive questions. Although both correlations were insignificant, these results indicate a conscientious respondent is more willing not to answer at sensitive questions.

4. Respondents choose not to answer significantly more at sensitive questions compared to non-sensitive questions.

To find out more about privacy concerns and social desirability this test measures whether or not respondents are more prone to choose not to answer at sensitive questions than when asked a non-sensitive question. Results of survey 1 are the only results that are applicable for testing, because this survey includes sensitive questions as well as the option not to answer. The number of respondents that have participated in this survey is 58.

A paired difference test is used when comparing two different sets or groups. This tests whether or not the means differ significantly and thereby the null hypothesis stands for having no difference between the means, while the alternative hypothesis says there is. In fact it is a one way alternative hypothesis claiming one mean is significantly larger than the other.

Paired test is calculated by the following formula:

$$t = \frac{\bar{d}}{\sqrt{s^2 / n}}$$

D- bar stands for the difference of the means.

s^2 is the sample variance, n is the sample size.

Number of NA-options chosen at sensitive questions = 18

Number of NA-options chosen at non-sensitive questions = 143

Only 21 of the 156 questions were considered 'sensitive'. In relationship to the total amount of questions. The number of respondents (n) is 58, so the in order to find the proportion of times the no-answer-option was used, the following formula is used :

*p (sensitive)= No. of no-answer-options chosen / ((no. of sensitive question * n) - missing no. of sensitive questions answered)*

p(non-sensitive)= No. of no-answer-options chosen/ (no of non-sensitive questions - no. of non-sensitive questions answered) n) - missing no. of non-sensitive questions answered*

P(sensitive questions = 18 / (1218 - 249) = 0,01857

P(non-sensitive questions) = 143 / (7830 -1521) = 0,02266

Test results:

P-value: 0.439545

Test statistic: 0.153433

No significant evidence against the null hypothesis is found, therefore the null hypothesis is not rejected. Remarkable is that the proportion no-answer-option at non-sensitive questions was higher than the proportion at sensitive questions. The expectation based on the literature review was a higher rate of no-answer-options chosen at sensitive questions. This result indicates there it attention or social desirability to respectively increase the number of no-answer-options chosen at non-sensitive questions or lower amounts at sensitive questions.

5a) Respondents report significantly more social desirable answers without the no-answer-option

Sensitive questions have two types of answers; social desirable and undesirable. Choosing not to answer or a neutral answer at sensitive questions shows either lack of motivation or the respondent is concerned about his privacy. If the respondent is concerned about privacy and he probably has something to hide, he can answer desirably or choose the no-answer-option. Because of this all other answers beside the social desirable ones are coded as social undesirable ('totally agree', 'agree', 'neither disagree nor agree' and 'no answer'). Desirable answers are coded a '1' and undesirable answers a '0'.

In other words surveys 1 and 3 include sensitive questions, that results in a total of 135 respondent. In order to find out if respondent give more social desirable answers than undesirable answers coding is divided into two groups:

Social desirable answers: 1

Social undesirable answers: 0

The independent t-test on the number of social desirable answers measures if this number is higher without the option not to answer. It compares the means between two unrelated groups. The expectation is that the no-answer-option leads to less social desirable answers($\mu 1 > \mu 2$), because this option gives respondents the feeling their privacy level is higher

SPSS provides us with the following data: (table 19 and 20):
Table 25 shows how the mean of group 1 (socially desirable answers without the no-answer-option) is larger than group 2, where the no-answer-options is included. These relations are conforming the expectations.

NA	Number of respondents (n)	Mean (no. of socially desirable answers)	Std. Deviation	Std. Error Mean
No-answer-option excluded	76	12,4342	5,96285	,68399
No-answer-option included	58	12,0690	5,86653	,77031

Table 19

Table 20 indicates the test results at the independent T-test, showing that there is no significant difference between the means of both groups.

Levene's Test for Equality of Variances						
F-value	Significance.	t-value	Degrees of freedom	Significance (2-tailed)	Mean Difference	Std. Error Difference
,007	,933	,354	132	,724	,36525	1,03243
		,355	123,818	,724	,36525	1,03016

Table 20

No significant difference between these two groups does not indicate a positive influence on privacy concerns by adding a no-answer-option to sensitive questions. While insignificant, the higher mean

of socially desirable answers given when the no-answer-option was excluded, encourages further research on the psychological effect of the no-answer -options.

b. Illegal sensitive questions have significantly higher rate of no-answer-option compared the legal sensitive questions

This hypothesis claims that people are more concerned with their privacy if they are asked to fill in a question about illegal activities. This is also a paired proportion test , in which the proportion of two groups are compared. The proportion of the number of no-answer-options chosen at illegal sensitive questions is expected to be higher than the proportion at sensitive questions concerning legal issues. Only the results of survey 1 are applicable for testing, because this survey includes both sensitive questions as the ability to choose the no-answer-option. The number of respondents that have started this questionnaire is 58.

Number of NA-options chosen at illegal sensitive questions = 10
Number of NA-options chosen at legal sensitive questions = 11

Only 21 of the 156 questions were considered 'sensitive'. Ten questions were about illegal activities and eleven about legal activities.

*p (illegal)= No. of no-answer-options chosen at illegal sensitive questions / ((no. of illegal sensitive question * n) - no. of illegal sensitive questions unanswered)*

*p(legal)= No. of no-answer-options chosen/ (no. of legal non-sensitive questions * n) - no. of legal non-sensitive questions unanswered)*

P(illegal) = 10/ (638 - 104) = 0,01872
P(legal) = 11/ (580 - 128) = 0,02433

Test statistics: 0.2059914
P-value: 0.419015

These measurement indicate no significant higher proportion that chose not to answer at legal sensitive questions compared to illegal sensitive questions. Remarkable is the fact that legal sensitive questions show a higher rate of no-answers than illegal sensitive. These results indicate further

research has to be done (chapter 6.3) on the degree of sensitivity, because it was expected that illegal sensitive questions would cause more no-answers than legal sensitive questions.

Chapter 6: Discussion

This chapter combines interpreting the results by test analysis (chapter 5) and comparative research with the existing theories that are a result of the theoretical framework; the literature review (chapter 2). The purpose is to test existing theories and add value to the discussion. In order to find out the added value to future research, limitations are to reckoned. Subsequently recommendations are given for exclusive continuation of this research as the theoretical framework contains some contradictions. Paragraph 6.1 is all about the results of the tests in relationship to the hypotheses, while the second paragraph of this chapter marks the conclusions made in this thesis by answering the research questions. At last recommendations are given for further research (chapter 6.3).

6.1 Interpreting results

Based on the data analysis a couple of unexpected results were presented and also some tests turned out to underline existing theories. This paragraph benchmarks the analysis, practical tests, with the theories noted in the literature review.

In the first hypothesis the effect between the provision of the taskbar and motivation was extensively tested resulting in a higher diversity of answers in the second part of the survey, including no significant positive effect when the taskbar is provided. These results contradict the finding done by Myers (2010) that a taskbar has positive influence on the motivation, but it does support the common agreement that elaboration does not change during the survey. As cognition is an individual trait according to Petty and Cacioppo (1986), respondents choose to take the central or peripheral route of processing the survey and this elaboration does not show a decrease over the length of the survey comparing the first and second half.

Remarkable on the data analysis of the second hypothesis was finding out that adding a taskbar does not have a significant higher completion rate. These findings contradict the notion (Conrad et al, 2010) that the taskbar giving progress feedback seems to have a positive impact on completion rates. Although this was the common agreement this result adds to the study done by Couper et al (2001) that claimed giving feedback about the progress in the survey would display no difference in completion rates. The second part of the hypothesis claimed no significant decrease between the amount of breakoffs in the second half by comparing the results between respondents with and

without provision of the taskbar. Important to take in account is the large difference between the amount of breakoffs per half; 90% of the breakoffs occurred in the first half. The fact that no decrease was found as a result of adding a taskbar is in line with the assumption made by Boltz (1993) that people allocate mental resources for a task based on how long they expect the task to be. In the invitation the number of minutes that were expected to complete the survey was mentioned and was realistic, so the number of breakoffs will decrease in the second half, because respondents are not feeling disappointed about the length of the survey.

The third hypothesis concerning conscientiousness and its relation to the number of times respondents choose not to answer is related to the need for cognition as discussed by Petty and Cacoppio (1982) and their concerns about the extent of engagement and enjoyment of effortful cognitive activities. Conscientiousness is an indicator for motivation based on the five-factor-model of personality traits by Costa & McCrea (1992) and is a degree of attention based on exuberance and therefore conscientiousness is expected to have a significant negative effect on choosing not to answer. Data analysis prominently found no significant relation between these variables, although adding sensitive question resulted in a positive relation, while without sensitive questions a negative relation was found to the number of choices not to answer. This causes an insignificant interaction effect between sensitive questions in the relation between no-answer-options and motivation. These finding are contradicting the assumption made by McKenzie et al (2006) that the option not to choose is used earlier when stakes are small or that people do not fully compensate the effort of answering the survey question with the importance of making the decision. Although both correlations were insignificant, these results indicate a conscientious respondent is more willing not to answer at sensitive questions.

Testing whether sensitive questions cause respondents to choose significantly more no-answer-options than non-sensitive questions found no significant increase between both means. Based on Wish, Hoffman and Nemes (1997) who found misreporting on drug abuse, one of the sensitive topics that is used in the surveys, a significant difference could be expected. But this result is in agreement with Johnson and Goldstein (2003), who have concluded that choosing not to answer might occur caused by reference dependence, whereas this answer determines that strength or weakness of the answers on the following questions. The other important reason for this is social desirability.

Social desirability was tested in the fifth hypothesis and there was no significant higher amount of social desirable answers reported. This leaves reference dependence (Johnson & Goldstein, 2003) as

an explanation on sensitive questions reported no significant increase of the no-answer-option, this also should be examined in further research (chapter 6.3).

According to Joinson, Woodley and Reips (2007) respondents are more likely to answer sensitive questions with 'prefer not to state' if the respondents have the feeling that their anonymity is compromised, for example by questions on highly personal subjects. As the use of the 'I prefer not to state' as answer to a sensitive question is methodologically similar to the use of the no-answer-option according to Knapp and Kirk (2003). This implies no higher rate of times a respondent chose not to answer at illegal versus legal issues is expected. The last hypothesis and the results empathize these statement by having no significantly higher rate for sensitive illegal topics.

6.2 Conclusions

The conclusion is based on answering the research question : "How could privacy and motivation be enhanced through better survey design? ".
This study's purpose is to test existing theories and add value to the discussion on how the provision of the no-answer-option, taskbar and sensitive questions influence privacy and motivational concerns. Privacy and motivation are key elements to enhance data validity. Answering this question is done by following up the research questions (chapter 1.2). The first research question concerned finding out more about systematic response tendencies. A lot of research has already been done in this field of work, The common-method variance is the systematic response tendencies on which this study focuses. This kind of variances are caused by measurement errors that lead into misleading data results.
Assessing the effect of the no-answer-option, taskbar and sensitive questions on survey validity could be done in two ways: ex-post and ex-ante remedies. The choice for the latter is based on a couple of reasons. The first reason is that managerial relevance would definitely erode as managers are not going to take all very complex formulas every time data is provided by a survey. The other main reason is that prevention is a much better (financial) option than curing the data afterwards.

This study makes a distinction between the taskbar, no-answer-option and sensitive questions by linking the taskbar solely on the motivational concerns, while the no-answer-option with an interaction effect of the sensitive questions are united to both motivational and privacy concerns. There is a contradiction looking at academic literature about the relationship between the taskbar

and motivation of the respondent. Data analysis show that there is no significant positive relationship between giving feedback of the respondent's progress in the survey. Notable is that the diversion of answers is significantly larger in the second half of the survey indicating that the respondent kept changing a lot indicating motivation, although the same result is shown at the data of the surveys without the taskbar. Motivation can also be judged by completion rates and the number of breakoffs. The completion rates show no significant increases when the respondent gets feedback about the progress made in the questionnaire. Notable is the result that the number of breakoffs is significantly higher in the first half of the survey, but the provision of the taskbar appears to have no significant effect on the decrease of this number in the second half of the survey.

The no-answer-option was expected to have both an influence on motivation as privacy. The number of times a respondent choose not to answer is the result of disinterest, but also a sign for privacy concerns as the respondents does not want to answer a sensitive question for instance. These sensitive questions did not result in a significantly higher amount of times respondents use the no-answer-option, indicating social desirability plays a role when privacy concerns occur, but also the number of social desirable answers is not significantly higher when the no-answer-option is added to the Likert-scale.
Motivational concerns are measured using aforementioned methods (completion rates, diversity of answers), but conscientiousness is also a way to measure if a respondents finds him- or herself motivated as is explained in the five-factor-model of personality traits. Conscientiousness has a positive relation to choosing not to answer when sensitive questions were included to the survey. The relation was positive without the question, suggesting conscientiousness is a trait that lead to a higher amount of times the no-answer-option is chosen.

The analysis of the data show no great discrepancies with many existing theories, but they do add value to the discussion on how design factors enhance privacy and motivation. For instance, the tests show that the provision of a taskbar have no significant positive relationship to a higher completion rate, which contradicts many former tests on this subject. Another remarkable result was that the diversion of answers, another indicator of motivation, increased by comparing the results in the first and second half of the survey. Adding sensitive questions to the survey resulted in a positive relationship between no-answer-option and the completion rates, although this relationship was not significant. Sensitive questions is a topic of many forms and degrees. To find out more about certain distinctions the amount of times the no-answer-option were studied and test results showed no significant increase of no-answer-options used at the illegal issues compared to the questions

concerning legal issues. These findings contribute to the discussion, because social desirability is difficult to measure and the examination done by this research showed no higher levels of social desirable answers.

In conclusion, this paper studied various ways to improve motivation/attention and privacy concerns by adding a taskbar, no-answer-option and sensitive question. The paper also tested how different sensitive questions related to each other and how motivation is over the course of a survey. All relationships between the design factors, motivation and privacy are tested and a lot of further research is needed, but both unexpected as confirming results are presented. This paper adds value to the acknowledgments of earlier works and gives perspective of potential further research.

6.3 Limitations and recommendations

Conclusions of this study are certainly very valuable, but this study is also an important showcase to find out what kind of studies are needed to find out more about these design factors and the way the influence the survey results. Enhancing motivation and privacy through the provision of the taskbar and the no-answer-option proved to be difficult and possibly differs if the methodology changes. Therefore the recommendations made are mainly around adjusting the methodology to find out more about the phenomena under different circumstances. Following suggestions for further research are suggested as a result of the research in this paper:

- A study with different lengths of surveys to discover if the effect of the taskbar relates to the length of the survey. This study works with two different lengths, because of supplementing the sensitive questions. According to former academic research sensitive questions have an effect on completion rates, so probably the results may be different.

- Diversity of answers and the influence of the taskbar was based on the standard deviation of the respondent. No difference was made between a respondent that filled on all questions and one that stopped after the first screen (11 questions). Further research could be done weighting each standard deviation on the amount of questions each respondent has answered.

- An expected result was that diversity of answers (standard deviation) decreased the second half of the survey. This result could have been influenced by the fact that the majority of the quitting respondents were still present in the results. Deleting these breakoffs results in ignoring valuable data and because there are also a lot of dropouts after the first set of questions as the total questionnaire was split into seven different screens. You could not break off in the middle of a screen, all answered questions would not be saved. These respondents would also influence the standard deviation of the first half. In retrospect the survey should be made where there are an even amount of screens (to make a clear distinction between the first and second half) or all answered questions should be saved so no data is lost.

- One or two possible outliers have a large effect on the result of the amount of breakoffs in the second half of the survey. A larger sample lowers the strength of the outliers, so this study recommends to do more research on this subject by using larger samples.

- Data of this study shows whether or not and on what page respondents decided to stop, but no information is given about after which question the respondent choose to stop. More research is needed about the amount of respondents stopping the survey abruptly and directly after questions are too sensitive. It provides more knowledge about the relationship between sensitive question and abruptly ending the survey.

- People may be okay with answering one or two sensitive questions, but it is possible that after a number of sensitive questions the 'mood' of the respondent changes. Further research about the reaction to lots of sensitive questions is acquired. Need for cognition might increase as is the need for privacy, so probably more NA-options are chosen or more socially desirable answers are given.

- Create a survey with screens filled with only sensitive questions and screens with only non-sensitive questions in order to compare standard deviations (diversity of answers) and find out more about the effect sensitive questions have on need for cognition. This construction enhances the possibility to compare the diversity of answers between screens.

- The development of numbers of NA-options should also need to be examined in further research. Do respondents tend to give more or less 'no-answers' as the respondents run

down the survey? A follow up question is too examine if this happens because of privacy or cognitive concerns.

- Legal sensitive questions can also be divided by the degree of severity or a degree of how personal the question is. For instance whether or not a respondent has peed publicly is a lot less personal than asking whether or not a respondent wishes he did not have any family for a majority of the respondents. Using the degree as a tool to measure whether or not as the degree increases the number of times the respondents choose not answer increases. Creating groups of sensitive questions per subject (family e.g.) give more information about the different kind of sensitive questions and the relationship to attention and privacy concerns. Different questions have to added to the questionnaire to make such a division possible and also the subjects of the sensitive questions could be changed, for instance by adding questions about salary of political preference.

- Method variance is different per field of interest according to Cote and Buckley (1987), so by using different topics or one specific subject on which the survey questions are based, these results could differ a lot. Further research could be done on all hypothesis using a lot of different field of interests and compare results on every account.

- Design factors concerning the Likert-scale and fixed question have influence on completion rates and therefore on data concerning motivation. Changing to open questions or different scale variables influence motivational concerns as does the influence of for instance the taskbar. Diversity of answers will not be a measure of attention anymore, it changes the whole dynamic of research.

- The surveys are designed in a specific way to gradually go from one topic to another and without a lot of surprising twists or questions phrased both in a negative as positive way to keep the respondents sharp and acute. These are tools to find out more about the attention of the respondent if conflicting answers are shown. This method could also be used in further research as completion rates and diversity of answers gave no significant results.

This paragraph concludes that minor changes in the approach could be done to find out more about privacy and motivation. The way data is used and interpreted are also ways establish or conflict with the research done by this paper.

Reference list

Bagozzi, R. P., & Yi, Y. (1991). Multitrait–multimethod matrices in consumer research. *Journal of Consumer Research, 17,* 426–439.

Blumenthal, Mark (2008). "New Pew data on cell phones". Pollster. Retrieved 2013-06-04.

Boltz, M.G. (1993). Time estimation and expectancies. *Memory and Cognition (21) p.853–863.*

Bradlow, E.T. & Zaslasvsky A. M. (1997). A hierarchical latent variable model for ordinal data with 'no answer' responses. *Journal of the American Statistical Association, Vol. 94, March 1999.*

Cacioppo, John T.; Petty, Richard E.(1982), The need for cognition. *Journal of Personality and Social Psychology*, 42, 116-131.

Campbell, D. T., & Fiske, D. W. (1959). Convergent and discriminant validation by the multitrait multimethod matrix. *Psychological Bulletin, 56, 81-105.*

Chaiken, S. (1980). Heuristic Versus Systematic Information Processing and the Use of Source Versus Message Cues in Persuasion. *Journal of Personality & Social Psychology, 39(5), 752-766.*

Chang, S.J., van Witteloostuijn, A. & Eden, L. (2010). Common method variance in international business research. Journal of International Business Studies 41, 178–184.

Cohen, A.R., Stotland, E. & Wolfe, D.M., (1955). An Experimental Investigation of Need for Cognition. *Journal of Social Psychology, 51(2), 291–294*

Conrad, F.G., Couper M.P., Touranqeau, R. & Petytchev, A. (2010): The impact of progress indicators on task completion. Interact. doi:10.1016/j.intcom.2010.03.001

Costa, P.T.,Jr. & McCrae, R.R. (1992). Revised NEO Personality Inventory (NEO-PI-R) and NEO Five-Factor Inventory (NEO-FFI) manual. Odessa, *FL: Psychological Assessment Resources.*

Cote, J.A. & Buckley, R. (1987). Estimating trait, method and error variance: generalizing across 70 construct validation studies. *Journal of Marketing Research, 24, 315-318.*

Couper, M, Traugott, M. & Lamias M. (2001): Web survey design and administration. *Public Opinion Quarterly. University of Michigan (65):230–253.*

Craik, F. I. M., & Lockhart, R. S. (1972). Levels of processing: A framework for memory research. *Journal of Verbal Learning and Verbal behavior,* 11, 671-684.

Crawford, S., Couper, M.P. & Lamias, M. (2001) Web surveys: Perception of burden. *Social Science Computer Review. 2001;19:146–162*

Crowne, D., & Marlowe, D. (1964). *The approval motive: Studies in evaluative dependence.* New York: Wiley.

Dillman, D.A., Smyth, J.D., & Christian, L.M. (2009). Internet, mail, and mixed-mode surveys: The tailored design method. *Hoboken, NJ: Wiley*

Dinner, I., Johnson, E.J., Goldstein, D.G., & Liu, K. (2011). Partitioning Default Effects: Why People Choose Not to Choose. *Journal of Experimental Psychology.*

Dole, J.A. & Sinatra, G.M. (1998). Reconceptualizing Change in the Cognitive Construction of Knowledge. *Educational Psychologist, 33, (2-3), 109–128*

Eagly, A.H. & Chaiken, S. (1993). Process theories of attitude formation and change: The elaboration likelihood and heuristic-systematic models. *In A.H. Eagly & S. Chaiken, (Eds.), The psychology of attitudes. Orlando: Harcourt Brace: pp. 303-350.*

Epstein, J. F., Barker, P.R., & Kroutil (2001). Mode effects in self-reported mental health data. *Public Opinion Quarterly, 65, p. 529-549.*

Fujii, E. T., Hennessy, M., & Mak, J. (1985). An evaluation of the validity and reliability of survey response data on household electricity conservation. *Evaluation Review, 9, 93–104.*

Harvey, L. (1987). Factors affecting response rates to mailed questionnaires: A comprehensive literature review. *Journal of the Market Research Society, 29, 341-353.*

Idaszak, J., & Drasgow, F. (1987). A revision of the job diagnostic survey: Elimination of a measurement artifact. *Journal of Applied Psychology,72,* 69–74.

Jacobs, D. (2013). Wat vinden leerkrachten van de onderwijshervorming ? *http://www.indemarge.be/survey-methodologie/wat-vinden-leerkrachten-van-de-onderwijshervorming/. Retrieved 2013-06-20.*

Jääskeläinen, A., Laihonen, H., Lönnqvist, A., Palvalin, M., Sillanpää, V., Pekkola, S. & Ukko, J. (2012). A contingency approach to performance measurement in service operations. *Measuring Business Excellence 16 (1),* pp. 43–52.

Jinks, A.M., Lawson V. &, Daniels R.A (2003). Survey of the health needs of hospital staff: implications for health care managers. *Centre for Health Policy & Practice, School of Health, Staffordshire University, Blackheath Lane, Stafford, UK.*

John, O. P., Naumann, L. P., & Soto, C. J. (2008). Paradigm Shift to the Integrative Big-Five Trait Taxonomy: History, Measurement, and Conceptual Issues. *In O. P. John, R. W. Robins, & L. A. Pervin (Eds.), Handbook of personality: Theory and research (pp. 114-158). New York, NY: Guilford Press*

Kahneman, D., Knetsch, J.L. & Thaler, R.H. (1991): Anomalies: The endowment effect, loss aversion and the status quo bias. *Journal of Economic Perspectives. 1991;5:263–291.*

Kahneman, D. & Tversky, A. (1972), Subjective probability: A judgment of representativeness, *Cognitive Psychology* **3**: 430–454

Kaplan, R. M., & Saccuzzo, D. P. (2009). Psychological testing: Principles, applications, and issues. *Belmont, Wadsworth.*

Leeuw, E.D. de (1992). Data quality in mail, telephone and face to face surveys. *Vrije Universiteit Amsterdam*

Martin, C. L., & Nagao, D. H. (1989). Some effects of computerized interviewing on job applicant responses. *Journal of Applied Psychology,74, 72–80.*

McKenzie, C. R. M., Liersch, M. J., & Finkelstein, S. R. (2006). Recommendations implicit in policy defaults. *Psychological Science, 17, 414–420.*

Myers, B. A. (1985). The importance of percent-done progress indicators for computer-human interfaces. *In Proceedings of the 1985 SIGCHI Conference on Human Factors in Computing Systems (San Francisco, California). CHI '85. ACM Press, New York, NY, pp. 11-17.*

Peterson, R. A. (2000). *Constructing effective questionnaires.* T*housand Oaks, CA: Sage.*

Petty, R. E., & Cacioppo, J. T. (1986). Communication and Persuasion: Central and Peripheral Routes to Attitude Change. *New York: Springer-Verlag.*

Podsakoff, P.M., MacKenzie, S.B., Lee, J.Y. & Podsakoff, N.P. (2003). Common method biases in behavioral research: a critical review of the literature and recommended remedies*. Journal of Applied Psychology 2003, vol. 88, no. 5, 897-903.*

Richardson, H.A., Simmering, M.J. & Sturman, M.C. (2009). A tale of three perspectives: Examining post hoc statistical techniques for detection and correction of common method variance. *Organizational Research Methods 12, 762–800.*

Salancik, G. R., & Pfeffer, J. (1977). An examination of the need–satisfaction models of job attitudes. *Administrative Science Quarterly,22, 427–456.*

Stulginskas, J.V., Verreault, R., Pless, I.B. (1985). A comparison of observed and reported restraint use by children and adults. *Accident Analysis & Prevention, 17, 381-386.*

Tourangeau, R. (2004). Survey research and societal change. *Annual Review of Psychology, 55(1), 775-802.*

Tourangeau, R., & Smith, T. W. (1996). Asking sensitive questions: Theimpact of data collection mode, question format, and question context. *Public Opinion Quarterly, 60, 275–304.*

Tourangeau, R., Rips, L. J., & Rasinski, K. (2000). The psychology of survey response. *Cambridge, England: Cambridge University Press.*

Tourangeau, R. & Yan, T. (2007). Sensitive questions in surveys. *Psychological Bulletin, 2007, Vol. 133, No. 5, 859–883.*

Wainer, H. & Kiely, G.L. (1987). Item clusters and computerized adaptive testing: A case for testlets. *Jounal of Educational Measurement, 24, 185–201.*

Williams, L.J., Cote, J.A. & Buckley, M.R.(1989). Lack of method variance in self-reported affect and perceptions at work: reality or artifact? *Journal of Applied Psychology, 74, 462-468*

Williams, L.J., Hartman, N., Cavazotte, F. (2010). Method variance and marker variables: A review and comprehensive CFA marker technique. *Organizational Research Methods 13, 477–514.*

Wish, E. D., Hoffman, J. A., & Nemes, S. (1997). The validity of self-reports of drug use at treatment admission and at follow-up: Comparisons with urinalysis and hair assays. *In L. Harrison & A. Hughes (Eds.), The validity of self-reported drug use: Improving the accuracy of survey estimates (200–225).*

Yan, T., Conrad, F.G., Couper, M.P & Tourangeau, R. (2011). Should I stay or should I go: The effects of progress feedback, promised task duration, and length of questionnaire on completing web surveys. *International Journal of Public Opinion Research*, 23: 131-147.

Yu, J. and H. Cooper (1983). A Quantitative Review of Research Design Effects on Response Rates to Questionnaires. *Journal of Marketing Research:36-44.*

Appendix

Appendix 1: All sorts of common method variance

General cause	Specific cause	Definition
Common rater effects	Consistency motif	Refers to the propensity for respondents to try to maintain consistency in their responses to questions.
"	Implicit theories	Refer to respondents' beliefs about the covariation among particular traits, behaviors, and/or outcomes.
"	Social desirability	Refers to the tendency of some people to respond to items more as a result of their social acceptability than their true feelings.
"	Leniency biases	Refer to the propensity for respondents to attribute socially desirable traits, attitudes, and/or behaviors to someone they know and like than to someone they dislike.
"	Acquiescence	Refer to the propensity for respondents to agree (or disagree) with questionnaire items independent of their content.
"	Mood state	Refers to the propensity of respondents to view themselves and the world around them in negative terms or the propensity of respondents to view themselves and the world around them in positive terms .
Item characteristics effects	Ambiquity	Refers to the fact that items that are ambiguous allow respondents to respond to them systematically using their own heuristic or respond to them randomly.
"	Social desirability	Refers to the fact that items may be written in such a way as to reflect more socially desirable attitudes, behaviors, or perceptions.
"	Common scale formats	Refer to artifactual covariation produced by the use of the same scale format (e.g., Likert scales, semantic differential scales, "faces" scales) on a questionnaire.
"	Common scale anchors	Refer to the repeated use of the same anchor points (e.g., extremely, always, never) on a questionnaire.

Item context effects	Priming effects	Refer to the fact that the positioning of the predictor (or criterion) variable on the questionnaire can make that variable more salient to the respondent and imply a causal relationship with other variables.
"	Embeddedness	Refers to the fact that neutral items embedded in the context of either positively or negatively worded items will take on the evaluative properties of those items.
"	Context-induced mood	Refers to when the first question (or set of questions) encountered on the questionnaire induces a mood for responding to the remainder of the questionnaire.
"	Scale length	Refers to the fact that if scales have fewer items, responses to previous items are more likely to be accessible in short-term memory and to be recalled when responding to other items.
Measurement context effects	Predictor and criterion (P&C) variables at same time	Refers to the fact that measures of different constructs measured at the same point in time may produce artifactual covariance independent of the content of the constructs themselves.
"	P&C variables in the same location	Refers to the fact that measures of different constructs measured in the same location may produce artifactual covariance independent of the content of the constructs themselves.
"	P&C variables by the same medium	Refers to the fact that measures of different constructs measured with the same medium may produce artifactual covariance independent of the content of the constructs themselves.

Appendix 2: random list of contacts that recieved an invitation to the survey link, letter'A':

A	**contact**
Aart Bunschoten	mail
Aart Tijhof	mail
Abraham Snoeij	facebook
Acacia Kooij	facebook
Ad Bernouw	mail
Adeni Nydi	facebook
Adri de Ruyter	mail

Albert van Leenen	mail
Aldo Luikens	mail
Alex de Werker	facebook
Alex Maliepaard	facebook
Alexander Korsman	facebook
Alexandra Barton	facebook
Alexandra de Raaij	facebook
Alfons Thiel	mail
Ali Agayev	mail
Alicia Riksman	mail
Aline van den Burg	facebook
Alke Wijn	mail
Allesandro Pangrazzi	facebook
Allison Middelbosh	facebook
Amanda Lucilla Janssen	mail
Andy Jansen	mail
Angela Hazelebach	mail
Angeline van der Aarsen	mail
Angelique Moerland	facebook
Angeloz da Longo	mail
Anita Sweers	mail
Anita van der Perk	mail
Annaleah Melsted	mail
Anne Kegge	mail
Anne van Dam	facebook
Anneke Gross	facebook
Annelieke de Hoop	facebook
Anne-Marie van den Hoek	mail
Annemiek Gommans	mail
Annette Tas	mail
Annika Aerts	mail
Anouk Lukkeren	mail
Anouska Bergmans	mail
Ans Bastiaanssen	mail
Anton Jonkman	mail
Antony peskens	mail
Arend-Jan de Leeuw van Weenen	mail
Arie Horstink	mail
Arie Monteny	mail
Arjen Verberk	facebook
Arlette de Lange	facebook
Armin de Rooij	mail
Aron Verboon	mail
Arthur Everaars	mail
Arthur Hamers	mail

Arthur Steijvers	facebook
Ashantie Eustace	mail
Astrid Eijsberg	facebook
Atze Vonk	mail

Appendix 3: the invitation

Thank you for participating in this survey that is obtained for the purpose of a Master of Science thesis of the Erasmus University of Rotterdam. We find your opinion of great importance. There are no right or wrong answers. The answers are kept anonymous and strictly confidential. The survey speaks for itself and will take between 20-25 minutes of your time.