COPENHAGEN INSTITUTE OF TECHNOLOGY
AALBORG UNIVERSITET

Social Influence Bias in e-Commerce:

Exploring the Role of Social Information

Master Thesis in Human Centered Informatics

By: Kathrine Lindskov Pedersen
Supervisor: Toine Bogers

June 2015

Student no.: 20132257 - Group no.: 82

Characters with spaces: 126.492 - Pages: 52.7

Abstract

For the past ten years the Internet position as a retail platform has become more predominant, which has generated a steady growth in the number of e-Commerce websites. This increase has not just affected the online retail sector, but also caused the traditional offline retailers to create new and alternative marketing strategies in order to keep up with the new consumption patterns. On the other hand, consumers are embracing the possibilities of e-Commerce and spending more and more money in these virtual stores. With this development also comes a new type of social influence, which appears in almost every corner of an e-Commerce website. And is this type of influence grounded in opinion information created by the consumers themselves, which we, with this Masters thesis set out to examine.

With this thesis we aimed to investigate how social influence impacts purchasing decisions made on online shopping websites. By looking into this field this we sought to gain insights into the significance of opinion information, founded in our PS and four RQ's. These questions firstly led us through a review of related literature on the subject, which supported our knowledge of the field of inquiry and made up a foundation for our inquiries made in the RQs. In order to build a solid experimental research we further made use of an array of theories and methods, which throughout Chapter 3, help us, construct a survey design for the research of the phenomenon.

After having tested our survey design using a pilot study group, we conducted the main study on a sample population, recruited online. This left us with an extensive amount of sample data, which was organized and analysed (Chapter 4) with the purpose of answering the PS and RQs.

The analysis of our data led to the main insight that there seemingly is a distinct divergence between the believed effect of social influence in e-Commerce, and the effect, which we detected through our research. Here we found that the respondents were not as highly influenced by the presence of opinion information as we had initial assumed. Our research exposed there this was in fact little to no connection between the intention to purchase or recommend a product solely based on opinion information.

Keywords: Social influence bias · Online shopping · e-Commerce · Opinion information

Table of Contents

1. Introduction

For the past decade the popularity and position of online shopping and e-Commerce has grown immensely. With it has come a long line of e-Commerce companies, who are rapidly becoming competitors, and some might say a threat, to the offline retail market. In Denmark alone the percentage of transactions made online versus the ones made offline, shifted an entire 5% from 2012 to 2014, giving e-Commerce a market share of 20% in 2014 (FDIH, 2015, p. 6) Even though 5% might sound like minor change, it meant an almost 20 billion DKK increase in total sales (FDIH, 2015, p. 6). And by the look of it, these numbers are merely displaying what is a drop in the massive ocean that is the global e-Commerce market. Further, the increase in both interest and popularity of e-Commerce, and particularly online shopping websites, is not only created and increased by the mere presence of the e-Commerce sites, but also with the help of social media.

As this new form of media has gained ground, and with an expectancy of continued growth (e-marketer, 2013), social media has also made an entrance in the world of e-Commerce. Here, companies incorporate features or plugins on their landing pages as well as on their specific product pages to enable the consumers to directly communicate, interact, and share companies' content or products. To a great extend, these social elements provide consumers the opportunity to give a, for the most part, uncensored, and honest opinion. It has somewhat become the norm that when a consumer is viewing a product he or she also has access to share it on the social media of their choice. In some cases a pay-per-like or pay-per-tweet can substitute for paying in usual currencies by simply clicking the respective social media button. Moreover, when having purchased a product, consumers can rate, review, and give their honest opinion for peers, companies or seller to see.

This form of social influence and virtual word of mouth is described by Qualman (2009, p. 1), however in his terminology the phenomenon is called World of Mouth and is described as being the connection which provides consumers with the opportunity of immediate interaction with peers, across borders, and timelines. This means that consumers from all around the world have a unique opportunity of sharing any feedback they may have on anything from products, consumer experience, service etc. More importantly, it is now easier for consumers to share their opinions and ratings with other consumers on any social platform. This information is gold for an e-Commerce website, but what about the things that they cannot measure? Companies are only able to gather information about their customers' behaviour when they shop, communicate about it on social networks, or in ratings and reviews on the companies' own site. What is harder to measure is how this information affects other consumers and in their decision making process.

However, one of the key factors in enabling an e-Commerce businesses to thrive as well as they do (FDIH, 2015) is for them to know as much about their customers as possible. This is where it becomes even more evident how important all information regarding the consumer's actions is for an online shopping site, and how the opinion information could prove to be one of the most essential information groups of them all. Information on how social influence affects the consumer's buying behaviour could further be a way of getting an advantage in terms of marketing for these sites. Ultimately, the insights and knowledge on how and if consumers weigh opinion information highly in a purchase situation, could help a shopping site increase sales and overall revenue. Further, this knowledge would most likely include a number of other valuable gains, as described by a major European e-Commerce company, in a statement that displays that the company's is already observing and testing the impact of this information:

> *[...] Recommendations or comments on product pages can be helpful in terms of lowering a return rate (e.g. if a manufacturer produces larger or smaller than usual sizes in articles). It is*

> *also being tested whether comments and shares increase sales for articles, meaning if users are more likely to buy a product if it has high ratings or was recommended by users connected to them. (Company Interview, 2015, Appendix 1)*

Social influence was previously what could can be described as being local (Meyrowitz, 2006, p. 21), and was based on what consumers saw their friends, peers or someone of a higher social status do. Later, with the introduction of written media, and even radio, and TV, the sources of influences spread to being national or even global. Now, with the Internet playing a large role in society, a modern consumer is constantly exposed to the influence of peers, companies, and strangers, from anywhere. Within this in effect consumers have interpersonal relationships across the globe, and traditionally it is within these relations that social influence occurs as an effect of influence from one human to another or one group to another (Pickard, 2014). However, the consumer is not necessarily conscious of where this influence stems from, or even that their choices might be shaped by it. Throughout this thesis we will examine the different factors and conditions, which can influence consumers in their online shopping decisions. Further, we will investigate whether or not consumers rely on the statements, recommendations, ratings, and reviews, provided online by both friends and other consumers.

For comprehensive purposes we need to clarify that we will use the terms e-Commerce and online shopping extensively throughout the thesis and that these terms are used as interchangeable wordings. Our experimental design will incorporate two different websites, where one (Amazon) can be is described as a marketplace for multiple retailers and private sellers, and the other (Zalando) is what can be classified as a conventional business-to-consumer (B2C) online retailer. We find that e-Commerce covers all online commerce activities including the trade of both services and physical products. This entails that the term covers all forms everything from electronic transfers, mobile commerce and online marketplace, just to name of few. Online shopping is a form of e-Commerce, which covers direct retail sales to consumers.

1.1 Problem Statement and Research Questions

This section holds the thesis' problem statement as well as four research questions. The thesis investigates the tendencies of social influence bias in e-Commerce. The experimental design is conducted in order to uncover how and if consumers' choices when shopping online are influenced by different social variables. These variables can be, but are not limited to: reviews or ratings of specific products, similarity in product purchases made by peers, and recommendations made by peers, stranger, and well-known opinion makers.

Problem Statement

> *This thesis seeks to investigate how social influence impacts purchase decisions in online shopping.*

Research Questions

To investigate the problem statement, we defined four individual research questions. Each question will contribute to the examination of the phenomenon, and for the purpose of the data analysis the RQs will be extended with the use of a null hypothesis (H_0), as well an alternative hypothesis (H_A). These are added in Chapter 4., where we analyse and report on the raw data gathered via the experimental design. In the following we post the four different RQs along with an outline of its purpose.

The initial RQ posted is investigating how social influence is generally perceived, but also its specific role in e-Commerce. This will be executed by presenting the user a selection of different types of

social influence in the survey design (Subsection 3.2.2.2) and by reviewing related literature and research done on the phenomenon.

> **RQ1:** *How is social influence perceived in general and specifically in e-Commerce?*

In order to investigate social influence and its effect on purchasing decisions we further posted two RQs targeting this area more in-depth. Therefor the second RQ aims at investigating how different types of social influence can affect the consumer's purchase decision. The second RQ in this way inquires:

> **RQ2:** *How does the type of social influence affect the consumer's purchase decisions?*

The third RQ seeks to extend the investigation initiated in the second RQ by examining the, which impact source of influence could have on purchase decisions. Here we will look at the various types and if there is any difference in intention to purchase between these sources. The third RQ therefor inquires:

> **RQ3:** *How does the source of social influence affect the consumer's purchase decisions?*

Lastly our fourth and final RQ integrates the two preceding questions and investigates the interaction between source of influence and type of influence. By examining the interaction of the two variables against each other we hope to gain insights into the how customers perceive these two in different combinations. The fourth RQ inquires:

> **RQ4:** *How do the source and type of influence interact to affect the customer's purchase decisions?*

Together these four question form the basis for our inquire into the phenomenon of social influence and will further help us acquire a better understanding of the effects of this phenomenon's presence in an online shopping setting.. From here, the next sections will be used to describe the research methodology and theoretical foundations of the thesis. These are the methods and theories, which we will use to build our examination of the field, to build our experimental design and to report on our findings.

1.2 Research Methodology

This section explains how the experimental and survey design that was carried out in this Master Thesis was conducted, and how we used the chosen literature to support it. We also touch upon which paradigm within philosophy of science that our research is rooted in. Lastly, it must be noted that the use of literature for the research has been extensive, and not all sources are mentioned in this section, but merely reference in the methodology chapter.

First and foremost, it is important to establish that the research for this Master thesis can be divided into different two groups of research data. The primary research is undertaken in order to examine the phenomenon of social influence in e-Commerce. This research consists of an experimental research design in the form of a lab study, presenting the respondents with examples of social influence in a variety of e-Commerce settings. The experimental design is created using a method, inspired by the work Lazar, Feng, and Hochheiser (2010) and Kelly (2009). The survey will be distributed via email and social media websites like Reddit and Facebook, and serve as an examination of the

aforementioned PS. This is done in order to gain insights into the phenomenon at hand, and the effect that this has on consumers using modern e-Commerce websites, as described in RQ 1-4.

The methods used for the research are constructed from work by Baines and Chansarkar (2002) and Kelly (2009). The latter has also influenced the configuration of the pilot study, along with van Teijlingen and Hundley (2002). Further, Patton (1990) and Kuniavsky (2003) have contributed to the understanding of both sampling methods. The secondary research data is what we have labelled related work (Chapter 2), and the review of this is done within the field of social influence and look at different approaches to the examination of this field. The secondary research is done in order to understand how research on social influence has been done both in general, and specifically in a marketing setting. We will further use the reviewed work, as secondary theoretical literature, as we find much of this research as being highly relevant and important to the research we are performing here.

When investigating a phenomenon such as social influence, we must chose the research instrument, which best applies to this phenomenon. With the current research into social influence in online shopping, the human research instrument is the most apposite. Finally, we have chosen to define that the research performed in this thesis is based on the positivistic, or what we now would define as post-positivistic, paradigm (Pickard, 2013). This paradigm is building on the beliefs that a researcher working within its frames will be seeking "*...multiple perspectives from participants rather than a single reality.*" (Creswell, 2007, p.20). Adding to this the research founded in post-positivism is often using various levels of data analysis with the aid of software or computer programs. Moreover the perspective of this particular branch of philosophy of science builds on the concept that there will always exist some degree of uncertainty within a research process. Thereby, this paradigm is adding the discussion of probability and uncertainty to the way in which we view our research and its results. We see this paradigm as being a both relevant and solid foundation for the research conducted throughout this thesis.

1.3 Theoretical foundation

In this section we will name the theories, which were employed in the initial research to gain insight into the field where the phenomenon of social influence is found. These theories will also make up the foundation for both our understanding of the results discovered through our research and frame these in the analysis and discussion. Furthermore a collection of literature related to social influence will be reviewed later in this thesis, however in this section, the more comprehensive literature on the subject will be listed. The area of social influence is investigated with the work of Wren (1999) as a great source of knowledge. In his work, Wren presents four main terms within social influence; Obedience, Conformity, Independent and Collective Behaviour, and Leadership and Followers (1999). These four terms create the outline of his work, where various other theorists and researchers contribute to a nuanced and comprehensive look into the many aspects of social influence.

Further, we have looked at Tapscott's extensive research (2009) on how the first generation to grow up using digital media has contributed to the way in which online media, shopping, and interaction are continuously developing. This has provided both inspiration and insights into the changeability of this social aspect of the online landscape, and especially into online shopping. Tapscott also discusses fields such as group or mob-behaviour, herding, and gender studies, as being influenced majorly by the phenomenon (2009). The way social influence affects the aforementioned fields of research, lies closely to how we find social influence affecting online shopping, and can therefore be seen as interlaced. Lastly, we have worked with different marketing principles in the preparation and analysis of the results of our research. The marketing principles cannot be described as a particular theory, and we chose to work with an array of literature to obtain an extensive knowledge on the subject. Amongst

these is aforementioned Baines et al. (2002), Tullis and Albert (2008), and Chaffey, Ellis-Chadwick, Mayer, and Johnston (2009), who all deal with either traditional or online marketing research.

1.4 Motivation

There are two dominant reasons, which make up the motivational background for this Masters thesis. Here we seek to explain our incitement for the examination of social influence in e-Commerce, and what we see as the main objective for conducting this research: Initially we see the phenomenon of social influence as being a key area for marketers and e-Commerce companies. Everyday consumers everywhere are faced with a number of decisions, which can be important or insignificant, difficult or easy. This can range from what they wear, which route they take when going to work, what TV shows to watch, whether or not to buy organic produce, and which supermarket to spend their hard earned money in. All these choices are not ones, which most consumers spare a lot of thought, and at most would possible be inclined to believe they themselves are making these everyday decisions. And to a certain extend, the marketers already know a great deal about how to get consumers to respond to pre-set triggers, emotional stimuli etc. However, research on how consumers influence each other could provide an even more in-depth understanding of why and how consumers purchase as they do. Secondly, this thesis will seek to provide a contribution to the already existing body of research done by other academics or practitioners with interest of the phenomenon. The aim is to add to the extensive pool of knowledge on consumer behaviour, and to increase the attention given to social influence in e-Commerce. We hope to create an interest in, and shed light on the effects of social influence, and thereby push further research in the area.

1.5 Scope

This Master thesis is, as defined in the problem statement, focusing on examining the influence and effect of social signals and how this could lead to social influence bias in an e-Commerce setting. We have chosen to incorporate different website interfaces from the two e-Commerce companies Amazon and Zalando in the experimental design, and further use three independent variables: *Source of influence*, *Type of information*, and *Price*. The sources of influence are divided into two groups: *Customers* and *Friends*. These groups are then combined with one of three types of influence: *Recommendations*, *Ratings*, or *Reviews*. Lastly, two different price variables will be incorporated: *High* and *Low*. By choosing these specific variables, we simultaneously rule out a long list of variables, which have not be included in the study. An expanded list of these variables will be named in the Methodology chapter, where the reasoning behind the choices made will also be discussed in further detail. In order to test the different variables we will create a series of mock-ups that combine the aforementioned variables. Each mock-up will further contain a product from Amazon or Zalando, which are either high or low priced. The selected variables are examined using an experimental design in the form of a lab study incorporating the variables in a total of ten individual mock-ups.

Researching this field in greater extend could involve examining dependent variables such as click-rate and time spent on page, which could be measured using real-time experiments like eye or mouse tracking. Further, conversion and sales as an effect of social influence could also be measured by these methods or by incorporating A/B testing on an actual e-Commerce website. This method would further enable us to gain insights into how social influence effects dependent variables such as sales, click-rate and time spent on page. However, these variables are not part of the current thesis, and will therefore not be examined further. Other capacities that could be interesting to examine is, if and how consumers react when exposed to negative versus positive opinion information, or the effect of opinion information in different consumer groups such as experienced versus inexperienced online consumers, first time online shoppers versus frequent online shoppers etc. As this falls outside the scope of this paper, this will also not be further looked into. In the upcoming Chapter 2, we will

however look into the different ways of investigating social influence, as done by other scholars as well as researchers.

1.6 Relevance

In the previous sections, we described the foundations and scope of this thesis. In this section, we define which relevance the work done in this thesis will contribute with. Firstly, the aim for the thesis is to provide insights into if and how different types of opinion information on e-Commerce websites influence consumers. This research will further investigate if there is a dominant source of influence in terms of impacting the consumers and if so, which of the chosen sources of influence is the most dominant. The relevance of the research done in the thesis can be seen in connection to prior studies within the same field. However, while others have focused on areas such as social shopping, interpersonal relations etc., this thesis mainly focuses on the very basics of social influence bias between e-Commerce customers. This can be of great interest to marketers and e-Commerce companies, who could be able to profit greatly from gaining access to information about what triggers the consumers, what makes them choose one item over another, etc. With information like this at hand, it could possibly be easier to predict sales, promote and indorse products or brands via external influencers and opinion makers and leaders, and perhaps push individual products using positive opinion information. We see the experimental work done in this thesis as providing a good foundation for future research into the phenomenon of social influence bias.

In the upcoming chapter of Related Work, we will look further into which forms of social influence other researchers and scholars have examined in the past. This will provide a perspective as to how widespread the field of social influence is, and how social influence is a dominant factor in many different industries.

1.7 Outline

To round off this chapter, a summary of the remaining components that make up this Master thesis will be given. The second chapter, *Related Work*, will provide a review on various literature connected to the subject of social influence bias. From here we will move into the third chapter *Methodology*, which seeks to explain the background for the experimental design including experimental variables and mock-up design, as well as the survey design, including survey components and survey deployment. Finishing the *Methodology* chapter is a walkthrough of the how the pilot testing was conducted and which tools will be used in the data analysis. From here we will move onto the *Data Analysis* chapter, which features a preliminary discussion of the how the raw data will be divided and analysed and also incorporate an analysis and examination of the results of the experimental design. Lastly we will round off this Master thesis with a general overall discussion of the experiment and its results, followed by a conclusion, as well as a look into future opportunities and perspectives for research of the phenomenon. Figure 1 provides a visual presentation of the outline for thesis in its entity.

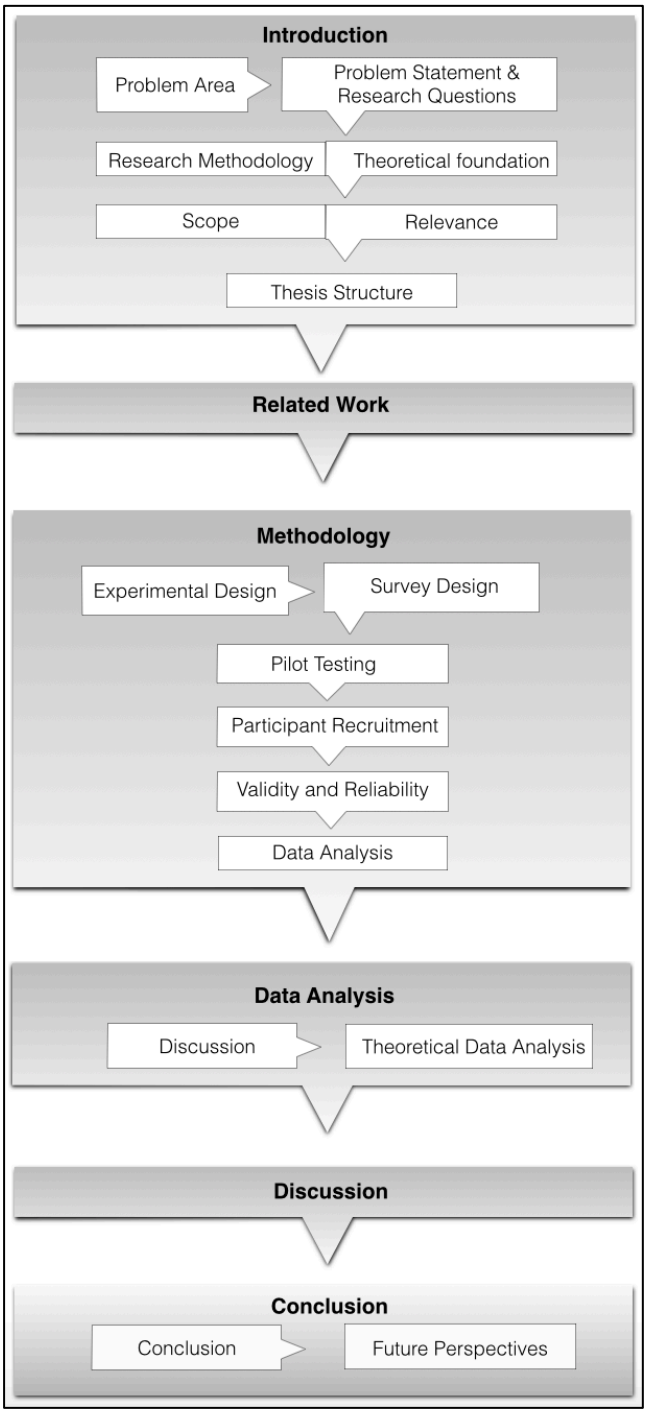


Figure 1: Structure of all Master Thesis elements

2. Related Work

Throughout this chapter a thorough review of several different scientific papers, reviews, related work and theories. which focus on the field of social influence. In the process of researching the chosen field of social influence bias, particularly in e-commerce, various sources concerned different aspects and themes related closely to social influence, have been examined. Throughout the following sections, some of the interesting and relevant literature will be discussed. The chosen reviews, articles, books etc. that all contain research and insight on social influence in correlation with areas such as marketing or gender studies. Further, most of the literature also touches upon themes within in the same field as social influence bias e.g. motivational influence, consumer purchase behaviour, social shopping, interuser relations etc. The overall aim of this review is to construct a better understanding of the field in which the current research will be done and the chapter will provide a wider picture of the multiple ways one could explore the phenomenon of social influence bias. The following literature contribute with findings that could be deemed interesting for this Master thesis.

More than 40 years ago, in 1974, Bonfield tested Dulany's 1968 theory of: "*...propositional control concerned with concept learning and verbal conditioning*" (p. 1), within a marketing setting. Bonfield tested Dulany's theory in the area of consumer brand purchase behaviour, in regards to both personal purchases, and purchases made by various consumer segments. Further, Bearden and Etzel's (1984) work, investigated the differences in consumer patterns when shopping for either necessity- or luxury products. Based on Bourne's matrix from 1957, the article looks into different reference groups and investigates their public and private consumer behaviour. Bearden and Etzel's article was followed by Childers and Rao (1992). Here they who look into further into how different reference groups influence their peers in terms of the various types of consumption behaviour.

Looking to Friedkin and Johnson (1999) an interesting discussion on social structure theory is found. Friedkin and Johnson's theory is centred on how a network is defined by its strengths and patterns. They argue that every social network has a dynamic, which ultimately could lead to opinion change for individuals in the group, as a result of group consensus and interpersonal agreement. Wang, Zhang, and Hann examine the same kind of group behaviour in 2014, which named this kind of behaviour *socially nudging*. The consumers' social behaviour prediction is the focus for Kim and Srivastava in their research paper from 2007. Moreover Kim and Srivastava look into how the consumers' online behaviour, e.g. recommendations, on social networks, can have impact buying decisions in these networks. Similarly, Hogg and Lerman (2014) in their experimental study, examines a user's reaction to social influence signals with offset Amazon's Mechanical Turk. Their aim is to understand how users respond to social signals and how this might be able to help predict social behaviour.

Yet another way of examining social influence is by looking at the difference between the genders. Two examples of the research done in this area is the work of Venkatesh and Morris (2000), and Carli (2001). The two publications look at two, somewhat different, perspectives of the gender related influences in our society. In their article, Venkatesh and Morris examine how the two genders are individually influence by different norms and perceptions when it comes to the decisions about and adaptation of, new technology. Carli's, however, constructs a more general review of how men and women are affected by different influencers in their choice of communicative channels. The findings made here points in the almost opposite direction of Venkatesh and Morris' research, which could lead to the assumption that gender indeed plays a significant role in social influence. However, the significance and influential factors is alternating in the context it is being examined.

Lastly, there are works, which have a more general focus on social influence. Both discuss the phenomenon on a larger scale, but do it in very different ways. The first of the two is Wren's (1999) publication on social influence that was previously mentioned in the Section 1.3. Four years later,, a different examination of the facets of social influence, was published by Cialdini and Goldstein (2003). In their article, they provide a detailed review of the development in social influence literature released between 1997 and 2002. Here they discuss the phenomenon in regards to aspects such as communication, affiliation, psychological influence and motivation.

The literature described in this chapter, has been reviewed in order to create a better understanding of the already existing work on social influence. This work covers many aspects, but it was especially the research of the phenomenon in relation to marketing and human sciences, which inspired provoked an interest in doing further research into the phenomenon of social influence. It must be mentioned that this literary review only covers a fraction of the published work on social influence and social influence bias, as the amount of the available literature on the subject far exceeds the extend of this review. However, we will use some of the supplementary literature will be used throughout thesis as secondary sources.

3. Methodology

Using the methodology described in the introductory chapter, the current chapter will discuss the decisions surrounding choice of methodology in further detail, i.e. explaining the basis for the chosen methodology and why the proposed experiment fits the area of inquiry. Moreover, this chapter's aim is to provide details into how the methods chosen will be used to investigate the phenomenon of social influence bias in e-Commerce. In the process of creating the experimental design for the purpose of researching this phenomenon many things have to be taken into consideration: It is of course important to remember that both *what* is being researched as well as *how* this is being researched, has a high impact on how the design is formed. This chapter serves as an overview of the entire process from brainstorming to the finished research design, outlining both obstacles as well as reviewing the supporting literature and theory.

3.1 Experimental design

The following subsections will discuss the details of our experimental design. First, we will outline the objective of the experimental design, followed by an review of our hypotheses, which is set up prior to the execution of the design. From here we will move on to an examination of the different dependent and independent variables, leading to the definition of the experimental conditions. The section will be rounded off by outlining how the interface mock-ups (Subsection 3.1.5), for the survey design were created with the use of aforementioned variables and conditions. All these elements will lead up to Section 3.2 that will address the survey design in its entity. We chose to work with this design method as it allows us to examine both the relationship between the variables in the interface, as well as the respondents' reaction patterns. Kelly (2009) explains the experimental design used in studies of interactive information retrieval (IIR), as being one that: "*...Examines the relationship between two or more systems or interfaces (independent variable) on some set of outcome measures (dependent variables)*"(p. 44). This is both in line with our chosen experimental- and survey design as well as our PS, with which we set out to investigate the effect of social influence in purchasing decisions in online shopping.

3.1.1 Objective

The objective of this experiment is to test the problem statement in a semi-controlled setting, thereby gaining insights into the phenomenon of social influence bias. We specifically aim to learn more about what motivates consumers to shop online, and if there is in fact a form of peer influence when it comes to making buying decisions. Moreover, we attempt to find a satisfactory answer to RQ 1-4.

3.1.2 Hypotheses

In order to explain the expectations for the experimental design and forthcoming survey, a hypothesis have been constructed for the process and execution of the experiment. We constructed a separate H_0 and H_A for analysis of the collected data. The $H_0$and H_A will be named in Chapter 4 in connection to each research question. However, our hypothesis for the experimental design is that the majority of online consumers who are influenced by the opinion information created directly on the e-Commerce website. We believe that online consumers are not solely buying on the basis of instinctive wants or needs, but also as an effect of what they see their immediate peers recommending or deprecating. However, in this experiment we can only compare the differences in types of influence. With further research and testing we would possibly be able to confirm or invalidate if social influence bias is in fact a phenomenon, which could be used in social marketing, and serve as an insight into customer behaviour.

3.1.3 Dependent and Independent Variables

The research undertaken in this thesis is constructed on the basis of a number of different independent and dependent variables. Examining these variables in an experimental design will enable us to manipulate the independent variables in order to test dependent variables. In the selection process in order to determine which variables to include in the experimental design, many different options were reviewed and discussed. This subsection will contain descriptive explanations and insight into the steps, and will provide overview of all variables. This will both include the ones selected for the experimental research and the ones, which we chose to leave out, however the primary focus on the first group.

The examination of which variables to include, and thereby investigated using the experimental design, was initiated by listing and discussing various dependent and independent variables that can be found on an online shopping website. With this we wanted to construct a wider frame containing everything that could be seen in relation to the phenomenon of social influence. From here we were able to narrow our selection further, in order to determine which variables fit the purpose of the experimental design and survey. Prior to the discussion and later selection of variables, it is important to establish the difference between dependent and independent variables, as the two are used in different ways. Here we use an example from Lazar et al. (2010) who explains the importance of defining the different variables in connection to the creation of a hypothesis: "*Independent variables refer to the factors that the researchers are interested in studying or the possible 'cause' of the change in the dependent variable….Dependent variables refer to the outcome or effect that the researcher are interested in*" (p. 25). In respect, independent variables are defined as a factor that we are able to change, and which are independent from the respondent's actions or opinions. On the other hand the dependent variable is what will show us the effect that the independent will have on these actions or opinions. Kelly explains this distinction as simply being the cause and the effect. "*Using the language of cause and effect, independent variables are the causes and dependent variables are the effects.*" (p. 37).

In this thesis we worked with two primary groups of independent variables, *Type of influence* and *Source of influence,* which were manipulated and combined in a number of ways in the survey design. Additionally, we will mention the use of quasi-independent variables, which are variables that are we as researchers cannot manipulate, but can create differences in the measured outcome (Kelly, 2009, p. 38). The selection of all variables within the mentioned groups will be discussed in the following subsections.

3.1.3.1 Source of Influence

When a consumer enters an online shopping site, they are not only met with information created with the purpose of informing about the products' properties, but they also see content created by other customers or even friends. The latter usually requires the consumer to be logged in on the website, and in most cases also to be connected to some sort of social networking platform (e.g. Facebook, Twitter, Google+, etc.). The user generated opinion information created by consumers or friends is what could be defined as 'real opinion information', as it comes from real consumers (Figure 2). On the other hand, it is necessary to establish that opinion information is often created by a different group, with the purpose of endorsing products on behalf of the e-Commerce website. We named this group of influencers *Authorities*, as we see them as having an immense conviction power, as a result of their already established name or brand. These types of product endorsements from celebrities and/or bloggers employed by online shopping sites can have an extensive impact on the consumer's purchasing decisions.

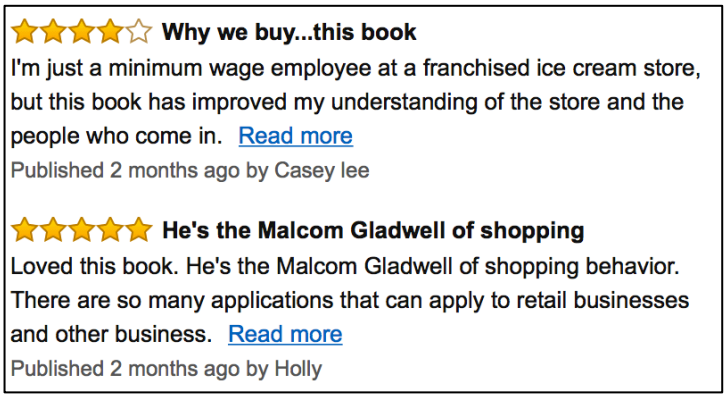


Figure 2: Opinion information provided by customers on Amazon (2015)

Further, so-called expert advice or recommendations from stylist or magazines can be some of the key influencers on e-Commerce websites. To this comes the third group of source of influence, which can be defined as being *Curated*. Here, the influence comes directly from the website itself, and here the opinion information being posted is typically in the form of offers on additional products, what they would recommend you buy, and information about savings. We will discuss which variables were chosen for the survey design in Subsection 3.1.3.4.

3.1.3.2 Type of Influence

From source of influence we moved on to look at the independent variable *Type of influence*. This group contains all the various types of opinion information that are presented on an online shopping site and which could possibly affect the consumer's decisions. This content is more or less featured on every product detail page, and on every online shopping site, and goes far beyond the simple like or favourite feature known from social media website. We will name all the types of information, which were discussed prior to the survey design.

The first group we address is the rating system that can be found on almost any type of online shopping website. Here, the consumers can typically give a product or service a rating in the form of stars, hearts, etc. The rating system can be displayed in a number of ways, either showing the average ratings for a product, the product's last x number of ratings, a distribution of all ratings (Figure 3), or even a summary of only positive or negative ratings. Here, the consumer often has the option of choosing how they want to have these ratings displayed. Furthermore, there are the text-based types of influence such as recommendations and reviews. The latter can be displayed in different formats such as snippets or whole reviews, and can in some cases even be voted up or down and/or answered. This is done when other customers find the review helpful or the opposite, and many times the exact number of votes are shown e.g. "*89 people found this review helpful*". Recommendations are quite similar to reviews and will usually displayed in full in close proximity to the product description. Recommendations further have the characteristic trade of being something positive, however some websites use them in conjunction with ratings and here a low rating will often be accompanied by a negative or explanatory comment. To label these kinds of comments as recommendations would be misleading, however we do see this occurring on e-Commerce websites like Zalando and therefore found it worth mentioning in this context.

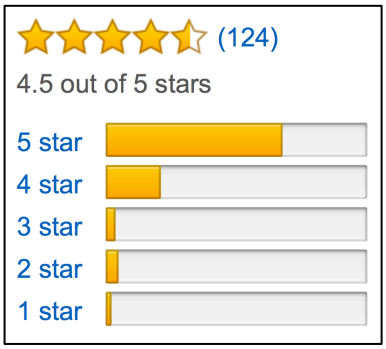


Figure 3: An example of rating distribution from Amazon.com (2015)

Moreover, we looked at the type of influence that are unconsciously created by the customers themselves, in the sense that the information is accumulated based on their browsing or shopping behaviour. However, these variables could just as well be created by the e-Commerce websites themselves, and therefore we have placed these in both the *Curated* and *Authorities* category (Table 1). These elements are presented with naming such as "*Other customers also bought:*" or "*People who viewed this also liked:*" usually with the intention to cross- and upsell. Furthermore, the elements, which display how many views a product has, the number of times it was shared via a social media, or informing that the product is on the bestseller lists, also fall under this last category.

3.1.3.3 Additional Variables

After having established the two main groups of independent variables, we also chose to look at a series of additional variables, which was discussed prior to creating the survey design. Initially, we discussed whether the variable of seasonal influences, such as the seasonable expiration date or timelessness of a product, would be significant to incorporate in the design. If a product is especially suitable for one season, e.g. swimwear for summer, but not for fitting for any of the other three seasons, can this influence the consumer's decision to purchase? This product will have a certain expiration date as an effect of the changing seasons. On the other hand, a product such as a pair of running shoes or a camera cannot be defined as a season product, which could define them as being timeless. However, we can also argue that newer and updated models eventually will replace this type of timeless products.

To this comes information such as price and availability that usually have a prominent place the product pages. If a product is priced high compared to what the consumer expect or what is within their budget, this could influence if the product is appealing to them or not. If the price is low, this could also create both negative and positive influence for the consumer. The negative reactions to a high or low priced product could be in form of mistrust e.g.: "*Why is the product so expensive/cheap compared to the offline retailer*". However, an opposite reaction could be created as an effect of a positive surprise: "*This product is normally out of my price range, but on this site it is much more reasonably priced*". When examining social influence we therefore have to consider that price is a highly relevant factor for most consumers, and that this will most likely have an impact on purchasing decisions.

3.1.3.4 Selection of Variables

In the previous subsections we discussed a number of different independent variables. To determine which variables would be the most relevant when researching the current phenomenon, it was necessary to look at all aspects and possible influencers in the field. In Subsection 3.1.3.2 we named the difference *Sources of influence* and ended up with four independent variables within this group: *Friends*, *Other Customers*, *Authorities*, and *Curated.* Subsection 3.1.3.3 outlined the different *Types of*

influence, and here we had the three main variable categories: *Reviews*, *Recommendations*, and *Rating*. As social influence in online shopping covers many different dependent and independent variables, an elimination process was undertaken in order to determine which of these variables were believed to be the most optimal for this particular experimental design. Following this selection process the two groups of independent variables remaining were combined to make the experimental conditions (Subsection 3.1.4) In the following we will elaborate this process and discuss the choices made in concerns to the final selection of variables.

In Table 1 an overview is provided, displaying how the different *Sources of influence* and *Types of influence* are combined to make the experimental conditions. All the independent variables are shown, however the variables, which have been crossed out, were not included in the final conditions (Subsection 3.1.3.4).

Source Type	Friends	Other customers	Curated	Authorities
Reviews	Snippet: Most helpful Snippet: Written by	Snippet: Most helpful Snippet: Written by		Snippet: Most helpful Snippet: Written by
Recommendations	Purchases Good together Also Bought Views Favourites Bestseller lists	Purchases Good together Also Bought Views Favourites Bestseller lists	Purchases Good together Also Bought Views Favourites Bestseller lists	Purchases Good together Also Bought Views Favourites Bestseller lists
Ratings	Average rating Rating distribution Last x ratings	Average rating Rating distribution Last x ratings		

Table 1: Independent variables discussed in Subsections 3.1.3.1 - 3.1.3.4

After the naming of all the possible variables, we were left with a long list of possibilities for the upcoming creation of the experimental conditions. As the aim was to create a design where we would be able to test our PS and RQs it was necessary to merely select the variables, which would allow us to achieve this aim. First, we selected which sources to include, and here we chose the two sources of influence *Other Customers* and *Friends*. The reasoning behind the choice was that *Curated* and *Authorities* did not contain the attribute of being social. They were viewed as having the characteristics of advertisement or promotional information created for the purpose of attracting the consumers to the specific products. This left us with the two consumers groups that were viewed as being sources of influence on a social level, as these groups were not seen as advocates for the retailer, but as merely as consumers.
Further, we had to decide which subcategories within the three main *Type of influence* groups should be included in the design. Our selection process here was based on primary two factors: The type had to be defined as social and further had to be connectable to the selected *Sources of influence*. On the basis of these two factors we further discussed which of the types would help assist us in creating the best possible survey design for the examination of our PS and RQs. We found that including variables from all three categories would be the most optimal when testing these, but only if these would apply to this examination of the PS and RQ. We chose *Reviews* written by *Other customers* or *Friends* over *Most helpful* by the same sources. This choice was made, as the respondent might not have any friends who shop online, and therefore displaying an element that is so specific could cause them to deem the scenario unlikely. Moreover we looked at *Recommendations*. Here, we made the decision to select only two variables from the group: *Purchases* and *Also Bought*. However, the latter was removed after

the pilot study (Subsection 3.1.4 + Section 3.3), being an up- or cross-selling instrument rather than being of influence of the purchase behaviour for the product in view.

The remaining variables discussed in Subsection 3.1.3.2 were not selected for the study as *Good together*, *Views*, *Favourites*, and *Bestseller lists* could not be defined as being of social character but rather automated information gathered to promote products. Also, the group *Ratings* was discussed and again we only chose the variable *Rating distribution*. *Average rating* would only provide a single numerical value for the respondent to consider, and *Last x ratings* would not necessarily provide a realistic image of all the ratings provided. Further, we considered the variables of price and decided to include both high and low priced products in the experimental conditions. With this, we would possibly be able to examine how this particular variable affects consumers' decisions. Lastly, we must name the quasi-Independent variables of *Age*, *Gender*, *Occupation*, and *Nationality*. These were not variables, which we as researchers would be able to influence, but they could however have an impact on the results. These variables were all applied to the demographics component and will be specified in Subsection 3.2.2.

3.1.4 Experimental Conditions

Through our research we wanted to examine whether and how the changes made in the independent variables, *Type of influence* and *Source of influence*, have an impact on the dependent variable, being the respondent's intent to purchase. Therefore we needed to combine these variables in a series of different experimental conditions, in order to manipulate and test them. Previously, we described the six specific independent variables that were chosen for the design and as we would be operating with two different e-Commerce websites, this number was doubled to make up a total of 12 variables; six for each of the websites. However, we also mentioned that this number was reduced after the pilot test of the survey design. Here we did not find the condition *Friends* or *Customers Who Bought This Also Bought* to be relevant for this study, as we did not see it having a clear social character as described earlier. Therefore these two versions with *Friends* and *Other customers* were removed from the surveys prior to the main study.

This meant that we were now operating with seven different variables, which had to be combined in order to produce the ten different conditions for the survey design. We had two different *Sources of influence* and three *Types of influence* and to this we should add the two defined variables of price. Now we had to combine the *Friends* or *Other customer* with a price, which was *High* or *Low* and finally add either *Review*, *Recommendations* or *Ratings*. In our selection process we had established that the variable *Friends* would not fit with *Reviews* as these would be in the form of multiple text-snippets. The probability of entering an online shopping site, choosing a random product and then seeing multiple *Reviews* written by friends, was simply very small. From this we constructed the ten combinations shown in Table 2., which made up our experimental conditions and the basis for our survey design. This design consisted of interface mock-ups each manipulated in order to contain the three independent variables of their specific experimental condition. The process of making the ten mock-ups will be described in the following subsection (3.1.5).

1. *Ratings* by *customers* on a *low* priced product	**6.** Recommendations by customers on a high priced product
2. *Reviews* by *customers* on a *low* priced product	**7.** *Recommendations* by *friends* on a *low* priced product
3. *Ratings* by *customers* on a *high* priced product	**8.** *Recommendations* by customers on a low priced product
4. *Reviews* by *customers* on a *high* priced product	**9.** Recommendation by *friends* on a *high* priced product
5. *Reviews* by *friends* on a *high* priced product	**10.** *Reviews* by *friends* on a *low priced* product

Table 2: Ten experimental condition constructed by the three groups of independent variables

The designs of the experimental conditions were all unique and we therefore could not test the conditions using a within-subject. We chose to work with a between-subject design, knowing that this could affect the outcome of our study. However, the choice was made as we would be working with a survey as our data collection method and that testing within-subject could cause the respondents to react to the change in the condition, instead of the change in the variable itself. Charness, Gneezy, and Kuhn explain this reaction, as being "*functions of circumstances*" (2011, p. 4) to the researcher demand.

3.1.5 Interface Mock-up Design

For the examination of the ten experimental conditions we constructed the same number of interface mock-ups. These were all based on product pages from the online shopping sites Zalando and Amazon. We chose to construct the mock-ups in the presentation software application Keynote, where screenshots of product pages were combined with pieces of opinion information posted to the site. In practice, this meant that the text snippets used for the *Reviews* and *Recommendations*, were actual opinion information that had been posted by real customers on these exact pages. The ratings were also based on ratings from actual products pages on the two website. The choice of using real *Ratings*, *Reviews*, and *Recommendations* was made in order to give the design a realistic look and wording. As mentioned previously in Section 1.5, we did nonetheless only select the snippets and ratings, which put the product in a positive light. The products for the designs were selected on basis of either their neutral appearance or their functionalities. This meant that we chose products like a TV, vacuum cleaner or camera, which had a wide appeal and would be attractive for a broad segment. On the other hand, we also chose products that would not attract a lot of smaller segments and could appeal to both sexes, like t-shirts, sneakers or shorts. Further, all items were chosen in neutral colours schemes and without an extra print or bold features.

Prior to assembling the different elements that would make up the interface mock-ups, we created two standardized wireframes: These simple wireframes, seen in Figure 4 and 5, were constructed to resemble the e-Commerce websites standard product interfaces. This was done in order to ensure that all mock-ups displaying products from either e-Commerce site, would resample the real product page, and have the website sites features and design. In this way, the respondents would not be exposed to a new design for each new condition they saw, but would be able to focus on the product and information. Further, it ensured that we as designers could standardize the placement of the opinion information.

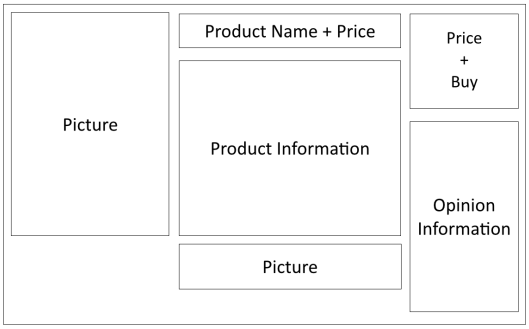


Figure 4: Wireframe for interface mock-ups (Amazon)

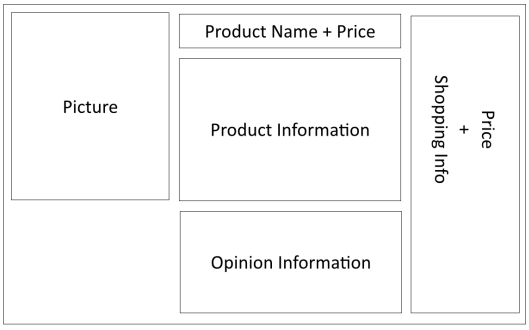


Figure 5: Wireframe for interface mock-ups (Zalando)

Finally, the standardized frame further ensured that elements, which were out of place or ones not holding an independent variable, could easily be discovered. These elements all contained additional information about things such as shipping, multiple prices for different versions of the product, as well as long descriptions. So, in order to both clean up the design and keep the standardized frame in all mock-ups these elements were removed. The final mock-up designs can all be viewed in Appendix 2.

3.2 Survey Design

In the following subsections the focus will be on how the six surveys were designed. Firstly, we will shortly outline our objective for creating the survey and which insight we expect to gain from it. From here we describe all the aspects of what our survey design entails. Further, we will discuss the importance of randomization and how this is implemented in our survey design, and lastly we outline the process of deployment of the survey design. With the study of social influence bias, and more specifically the occurrence of this phenomenon in e-Commerce, the experimental research is done via a qualitative method. This form allows us to obtain a comprehensive overview of the respondents approach to the phenomenon. A further, qualitative study as a research method also falls under the chosen philosophy of science paradigm with a structural approach, which could resemble a quantitative method (Creswell, 2007).

3.2.1 Objective

The purpose of the survey is to investigate the PS, and within the same frame aim at answering RQ 2-4. With this we are hoping to get a better understanding of the different variables' affect on consumers on e-Commerce websites. The survey may also produce insight into which influencers shown in the constructed survey conditions are the main incentive for the consumer's choices. These findings, strong or weak, will not be prime focus of the data analysis, however they could prove interesting in future research or discussions of the field and phenomenon.

3.2.2 Survey Components

As the construction of a survey design requires many different components, the following subsection (3.2.2.1 - 3.2.2.3) seeks to provide an overview of these. In here will be described each component and outline, its importance to the survey and the results. The three main components used for the survey design in this thesis were: *Shopping*, *Social Influence*, and *Demographics*. It was from these three main components that the survey was developed. Our survey design further entails the use of various kinds of question-formats, and the reasoning behind each choice will be outlined within the each subsection.

3.2.2.1 Shopping

As aim of the survey is to investigate the social influence phenomenon's effect in an online shopping setting, the component *Shopping* is of course highly significant. In order to investigate the nature of the sample population's online shopping habits, we initiated the survey by asking five questions relating to this topic. The questions in this component were closed, where two had the options of *Other* giving the respondent the option to add a category or answer himself or herself. The initial question was if the respondent ever would shop online or not, and had the options of answering *Yes*, *No*, or *I don't know*. If the answer here were negative or undetermined, the respondent would be sent directly to the questions on *Social Influence* (Subsection 3.2.2.2). If stating that they indeed did shop online they were sent directly to the second question on *Shopping*. These preceding questions were concerning the frequency and purpose of the online shopping, as well as what the respondents shopped for and the average amount spent (Appendix 3).

The second question, relating to frequency of online shopping, was posted as an fixed-multichotomous question (Baines et al. 2002, p. 100), which gives the respondents more than two possible answers to chose from. The respondents were however only able to provide one answer and they could chose between six different time intervals e.g. *1 - 3 times a month* or answer that they did not know the frequency. Following this, we asked the respondents about the most common reason there was for them to shop online. Here, the question was a multiple fixed-response (Baines et al., 2002, p. 101), where the answers implied a comparison to offline shopping, giving the respondents the

option to chose from four factors such as *Convenience* and *Better selection* (Appendix 3). Further, they could select *Other* and add another answer in the commentary box. Next, we inquired about the average spent when shopping online, and again we used a fixed-multichotomous question allowing them to give one answer with the option of providing an answer in the commentary box if deemed necessary.

The fifth and final question in the *Shopping* category was concerning to what products or services the respondent would shop for online. This question was also multi fixed-response and here we provided 11 different categories including a commentary box. This was the only question in the survey where more than one option could be selected. We selected this as one single respondent could easily be purchasing both groceries, electronics and clothing online. The categories for this last question were adapted from FDIH's analysis of the Danish e-Commerce (FDIH, 2015).

3.2.2.2 Social Influence

The social influence component is the main one of the three and was investigated using the earlier mentioned interface mock-ups (Subsection 3.1.5). To examine how the respondents were affected by the independent variables we exposed them to five, randomly assigned, experimental conditions. It was important to gather insights into which influencers were provoking an effect in the respondent's intention to for example purchase or recommend a product. Prior to presenting each mock-up to the respondents, we asked them to review and consider all the available information shown on the mock-up. This information was both textual descriptions of the product and price, as well as opinion information and of course the product itself. This was done in order to make the respondents aware that all this information could be important when having to answer the preceding question.

After having reviewed the mock-up the respondent was asked then to consider to what extent, if any, they had an intention to *Purchase*, *Save for future consideration*, *Recommend*, or *Share the product on social media or elsewhere*. Each of the options for the answers was to be provided by the respondent rating the intent on a five-point Likert scale. The information provided here would later form the basis for our analysis and could therefore be considered the core questions in our survey design. When having provided their answers to this question, the respondents were further inquired about the reasoning for providing exactly these answers. Here we chose to apply a multiple fixed-response question, but only allowing one response to each, as we found it sufficient to be informed of the main influencer that had affected their choice (Figure 6 + Appendix 3). The options provided here would help us to identify if the respondent simply liked or disliked the product in question, or if there was a more profound or specific reason for their answers. The options were all based on the information provided on the mock-up, such as *Product*, *Price or Value*, *Brand* etc. For these questions there further the option of adding an additional option in a commentary box.

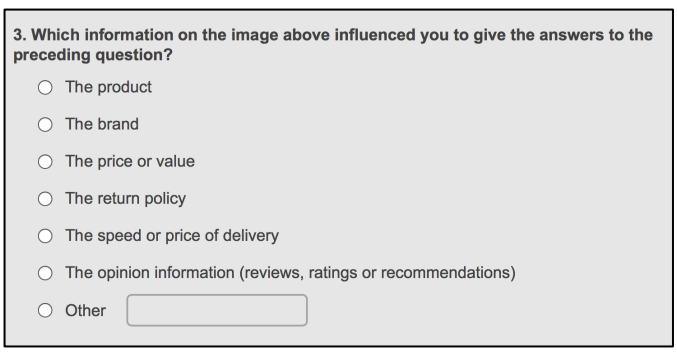


Figure 6: Social influence question

3.2.2.3 Demographics

The demographic data in this survey is collected in order to be able to investigate differences in behaviour between the various age groups, genders, and nationalities. This data will not be analysed as in-depth as the other components in the study. However, we did find it important to collect demographic data about the respondents, as this could prove to be one of the factors where we could be able to experience significant fluctuations in results.

All demographic questions were placed at the end of the survey. This choice was made based on the discussion on fatigue and boredom (Section 2.2.3). We did not want to risk respondents skipping the most important part of the survey, the *social influence* components, as an effect of having to answer demographic questions at the beginning of the survey. It was also estimated that even if respondents were to leave the survey after only answering the questions on *shopping* and *social influence*, this would still provide valid data. If we placed both the *shopping* and the *demographic* components before the *social influence*, the risk of not collecting enough valuable data do to boredom and thereby an increased drop out rate, would be significantly larger (Kuniavsky, 2003) The demographic questions were concerning the respondent's gender, followed by a question about his or her occupation; whether he/she was *employed*, *out of work*, *a student*, *retired* etc. (Appendix 3). Further, we asked about age, using intervals ranging from *Younger than 18 years old* to *75 years or older* and finally about *nationality*. As mentioned earlier, the demographical data collected will not be examined in the forthcoming analysis, however it is interesting as different patterns in online shopping behaviour might occur solely based on where the respondents are from, if they are male or female or how old they are.

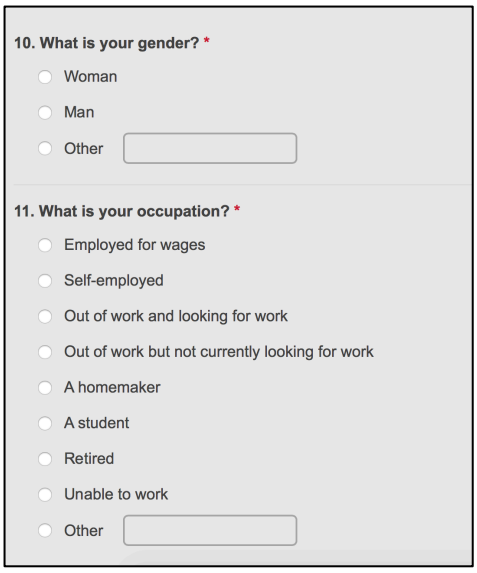


Figure 7: Demographic question

3.2.3 Randomization

When conducting a survey, the risk of respondents becoming bias is ever-present. Many things can cause this, but most commonly issues with fatigue, boredom or learning, is known to influence the responses and in the worst-case scenario; create invalid results (Kelly, 2009). One of the ways in which researchers can try to avoid bias and faulty results in surveys, is by using counterbalancing and randomization. In a total randomized experiment, it would not be possible for anyone, not even the researcher to predict, which experimental condition were shown to which respondent (Lazar et al., 2010). Opposite a basic design, as described by Kelly (2009, p. 51), would cause all respondents to be

exposed to the exact same conditions, in the same order. If we had been able to set up the first form, total randomization, the five experimental conditions would be presented in a randomly assigned order to every unique respondent. However, we were not able to achieve this automatic assignment, as the chosen account in SurveyGizmo (2015), did not provide access to such a feature. This meant that it was necessary to create the randomization in a different way.

We chose to create six different survey variants, as it would allow us to rotate the ten experimental conditions, so that they would appear in three variants, but the conditions could never be placed in the same order or more than once in a survey. We selected to do a round-robin assignment for randomizing of the experimental conditions and the result of this assignment can be seen in Table 3.

1	Ra-Cus-Low	Rec-Fri-Low	Ra-Cus-High	Rec-Fri-High	Rev-Fri-High
2	Rev-Cus-Low	Rec-Cus-Low	Rev-Cus-High	Rev-Fri-Low	Rec-Cus-High
3	Ra-Cus-High	Rec-Fri-High	Rev-Fri-High	Ra-Cus-Low	Rec-Fri-Low
4	Rev-Cus-High	Rev-Fri-Low	Rec-Cus-High	Rev-Cus-Low	Rec-Cus-Low
5	Rev-Fri-High	Ra-Cus-Low	Rec-Fri-Low	Ra-Cus-High	Rec-Fri-High
6	Rec-Cus-High	Rev-Cus-Low	Rec-Cus-Low	Rev-Cus-High	Rev-Fri-Low

Table 3: The six survey variants randomized using a round-robin method

To obtain an equally distributed amount of data for all the conditions these had to be tested on an equally sized sample population. As it was not possible to insured that this rule was adhered to completely, as we could not obtain a total randomization, we had to use a form of pseudo-randomization when distributing the survey variants as described in Subsection 3.4.2. One way of securing this pseudo-randomization was to make an entry in a group on a social network sites such as Facebook (2015) and ask possibly respondents to 'like' the post, whereafter we would send them a private message containing one of the six variants. Moreover, all variants were posted to the same subreddit on Reddit (2015) in order to maintain an equally distributed response rate.

3.2.4 Survey deployment

Using the Internet as a research instrument was an obvious choice as we are investigating an online phenomenon and needed a sample population that consisted of online-users and/or consumers (Section 3.5). Further, factors like speed, low cost, flexibility, and global reach will enable us to collect a larger amount of data, from a larger sample population in a shorter time without having to use more than one physical location and one computer (Baines et al., 2002). However, we did choose to use multiple platforms, in order to collect a more diverse and non-homogeneous sample.

The survey was created using SurveyGizmo (2015), an online survey tool favoured by many large corporations and companies e.g. the Danish Broadcasting Corporation (DR). SurveyGizmo offers a free trial period of 30 days, and in this the basic functionalities are available. The free version was used for the creation of the survey and following this a basic account was acquired in order to get access to more functionalities, as well as being able to access the raw data. The basis account did however not provide the necessary features to create the required level of randomization needed for the survey in question, as described in 2.2.3. It was on these grounds that we decided to create six different surveys. This solution was of course not optimal when having to handle the larger amounts of raw data from the survey. An upgrade to the Professional or Premier account would be needed to be able to gain access to the required functionalities e.g. randomization, however this was not possible due to the high costs of the upgrade, which meant applying a different method for randomizing the conditions, as described in Subsection 3.2.3. The surveys were all constructed in the same manner,

and with the same set of questions. For the randomization, to shorten the length of the survey, and to prevent fatigue, only five mock-ups were assigned to each survey variant. There are no rules as such to how long or how short the timing of a survey should be (Kelly, 2009), however the issue of fatigue was important to prevent as this could cause respondents to drop out of the survey prematurely.

3.3 Pilot Testing

A large part of studying a phenomenon like social influence is collecting knowledge and data concerning how *the public* views it. Social influence bias in e-Commerce can be an entangled field to investigate, as the reasoning behind the consumer's choices in the buying decision can be difficult to measure. Further, any survey design used for this type of research, must present a good structure, comprehensible language, and strive to prevent issues like fatigue. By conducting a pilot study on a smaller sample population we aimed at eliminating the majority of the design problems, pitfalls, and comprehension issues there could be in the main study. The pilot study is of great importance in the preliminary stages of an experimental study, and provides the researcher with a great amount of information and insight into both structure, interface and question design. Kelly explains the significance of conducting a pilot study:

> ...help researchers identify problems with instruments, instructions, and protocols; allow systems to be exercised in the same way they will be in the actual study; provide researchers with an opportunity to get detailed feedback from test subjects about the method; help researchers gain comfort with administering the study; and finally, they can be used to train inexperienced researchers. (2009, p. 60)

Thus pilot studies are not only conducted in order to rule out some of the plausible pitfalls and provide feedback from test-respondents, but also aim to ensure a higher success rate in the main study, explained further by Kelly "*...pilot tests help researchers identify and eliminate potential confounds and errors that might otherwise compromise the integrity of the study results"* (2009, p. 60). In this section, we will focus on how this study's pilot test was conducted, what feedback was given from the respondents, and how this information was used to optimize the survey for main study.

Our case the pilot study was conducted on a smaller group of only ten respondents over the course of two weeks. The number of pilot test respondents equals a sample size of just below 10% of the 200 respondents estimated for the participation in the main study. The sample size estimate is based on how many respondents it is likely for us to gather within the time frame for execution of the experimental design. This sample is not sufficiently large enough to provide us with in-depth insight of the phenomenon, however it will generate a wide and general insight in to the phenomenon at hand. Collecting in-depth information or data from a smaller sample population is described by Patton as being: "*...very valuable, especially if the cases are information-rich. Less depth from a larger number of people can be especially helpful in exploring a phenomenon and trying to document diversity or understand variation.*" (1990, p. 184). In the following paragraph we will review and discuss the use comments provided by five of the respondents. Some of these are translated to English from either German or Danish and these are marked with either DK or DE, and the original comments can be viewed in Appendix 4.

The first concern we saw was highlighting multiple times by one respondent. This concern was in regards to the look and placement of the mock-ups. These had been created using Keynote (Subsection 3.1.5), and directly placed in the survey. However, using Keynote meant that the standard size of the images were large when directly saved on the desktop. In the survey, this caused the images to take up more space than intended, that they were misplaced, could not be viewed in full extend, or that the question under the image was not aligned: "*Image/text alignment is off. Text*

content should be below image" (Appendix 4). This was also the case when opening the survey on a tablet or phone as remarked by one respondent: "*It does not work on iPad*" (Appendix 4).

Further the wording and phrasing of the questions was revised after reviewing the comments from the pilot study. Here the respondents noted that some questions and designations possibly would work better if posted in Danish or be revised in order to option a better overall comprehension. One respondent also noted that it was difficult to answer the question concerning intention to purchase, recommend, save or share, as the products were not the ones that he would ever want to purchase.
All of the comments and remarks were highly useful as it provided us with an insight into how the respondents would interpret the survey. This helped us improve the survey and conduct an overall revision of the design. This included a review of usability, change in wording or re-phrasing questions, display of images, and adding additional text in order to improve the understanding of the tasks. This process was important to get an understanding of how future respondents would react to the survey.

3.4 Recruitment of Respondents

When recruiting respondents for the study, we chose to use of various online platforms, to reach a much broader range of segments, as opposed to recruiting via just a single platform or channel. We chose this kind of probability sampling (Kelly, 2009) as one method of preventing biased results and to gather a sample, which was representative of the general population. It is of course not completely inevitable to prevent bias and errors, and we will discuss this and other issues further in Section 3.6. The different platforms used for recruiting respondents were Facebook, Twitter, and Reddit, where different generalized or personalized messages were used when approaching smaller sub-groups or individuals. Further, a small international group was approached via both Facebook and email by a third-party located in Berlin.

3.4.1 Segmentation of respondents

In order to recruit prospective respondents an identification of the primary population group was needed (Tull and Hawkins, in Baines et al., 2002). The primary population group in this instance was all people with access to the Internet, and thereby to online shopping. This is a rather larger primary population, which is why we needed to further set a frame for the smaller, sample population. As this experiment is investigating a phenomenon and further is focused on how this phenomenon influences consumers in general, only a few requirements were set. In order to create the frame and outline the more common characteristics of the demographic group, we looked to McKay, Pearson, Peart, Utreras, and Wang (2011). Their work describes a specific segment in the American population, named Generation X. Within this group, which by McKay et al. is set to be between 34 and 45, an entire 66% have previously purchased a product online (2011, p. 12). Further, this demographic group, as well as both older and younger are mainly buying products such as apparel and shoes, movies, music, and travels (2011, p. 13), when shopping online, and the same applies for the Danish consumers (FDIH, 2015, p. 11). Moreover, all age groups described by McKay et al., are relatively keen users of social media such as Facebook, Twitter, YouTube etc. So in terms of age, we find it necessary to target the survey to as wide a group as possible.

Looking at other demographic traits, we find factors such as nationality or geographical location to be interesting, but not highly significant for the research done in the study. If examining social influence and e-Commerce in relation to how the different markets react, this could be an interesting feature, so we chose to include this in the survey. This is further supported by the fact that getting a wider array of age groups, respondents from different educational background, different nationalities etc., could help preventing the study from being biased as a result of a too homogeneous respondent group, as

mentioned earlier. Therefore the only fixed guidelines for the demographics of the respondent groups were as follows:

- Respondents need to have access to the internet as this is where the survey will be distributed
- Respondents cannot all be from the same social group (e.g. same university or same workplace)
- Respondent must preferably be between the age of 18 and 75
- Both female and male respondents are needed

Lastly, we chose to place the demographic questions at the end of the survey. As explained earlier, this was done in an effort to not cause fatigue or even irritation as an effect of having to answer too many personal and off topic questions.

3.4.2 Global and Local Recruitment

The groups on Facebook, which was targeted, were all forums where the researcher either had a contributing or administrative role. Each of the groups can be seen as closed forum that is concentrated around a specific subject or interest, e.g. a group of university students or music fans etc. We needed to secure that these groups, who could produce a somewhat homogenous response pattern based on their similar background or interest, would not answer the same survey. Describing the purpose of the survey to the group members in the post and asking them to either "like" or comment did this, those who responded would then receive a personalized message containing one of the six survey variants for them to answer. In this way, we were able to control that a supposed homogenous group would receive and responded to different variants.

The use of Facebook and Twitter were the initial choices as six surveys were all distributed online to a range different platforms and social networking sites, in order to get a wide variety of respondents. First of all, the surveys were posted to the social networking site reddit.com, in specific to the so-called subreddit "Sample-Size" with 29,949 participants (May, 2015). The surveys were also posted to the subreddits "Favors" that has 30,505 participants and "Assistance" with 26,190 participants (May, 2015) in an attempt to recruit respondents; unfortunately none of these gave any responses. While trying to ensure the randomization and prevent bias, as described in Section 3.2.3, another third-party was used for the recruitment. Using email sent out and social media posts created by this third-party with an international network in Berlin, helped to provide a wide range of different respondents. With this way of recruitment, we further tried to ensure that the surveys were not only sent to a single homogenous group of respondents. Distributing the survey to several groups with a wide range of demographic properties including various nationalities and geographical location is important, as a too homogeneous respondent group can create an invalid sample (Kelly, 2009). The issues concerning validity will be further discussed in the following Section 3.5.

3.5 Validity and Reliability

The design of a survey and its components are crucial as to how the results of the study will be. This is why, when conducting any type of survey or study, we have to take into consideration that the results can be biased by a number of different factors. These can both occur in the survey itself or be caused by external influence affecting the respondents. Here, we look to Kuniavsky (2003), who explains why especially online or web-based surveys can be particularly vulnerable:

> "...because, lacking any direct contact with respondents themselves, they depend on the perceptions people have of themselves and their ability and willingness to accurately report those perceptions….They can only tell you what they think." (p. 304).

First and foremost, the validity of our survey results can be affected by the sample population's lack of completeness, meaning that the respondents stems from a too homogeneous sample group (Baines et al., 2002). In regards to the respondents, we also find that there is a potential risk of certain dropout rate. When an experimental design is being tested in a non-supervised lab study, the risk of respondents dropping out and thus not completing the survey can be increased. This can be due to the fact that the respondents do not feel a strong obligation to complete the survey. The drop-out can further be caused by respondents misunderstanding or not being able to interpret the questions or instructions giving in the survey (Baines et al., 2002).

As described earlier in Subsections 3.2.3 and 3.4.2, we applied randomization in order to obtain heterogeneous results in our sample data. The use of a Likert scale can be another factor, which can cause biased results. If the respondents get bored or tired of answering the survey, there is a possibility of provoking middle position bias (Baines and and Chansarkar, 2002, p. 106). Lastly, we must consider the risk of error and design flaws when analysing and reporting on the results of our survey. Here, a large number of hypotheses will be tested, which means that we will also work with several combinations of data. In worst case, having too many hypothesis, and thereby too many different factors to control can lead to errors that can affect the results and which might go unnoticed (Lazar et al., 2010).

3.6 Data Preparation and Analysis

In the fourth chapter, we analyse and report on the results gathered via the experimental design. The first step is combining all the raw data to create a complete overview. For this process, we used Microsoft Excel, which was chosen as it provides various methods for data processing, such as pivot tables and filter- and sorting options. Further, the program was especially useful as we worked with smaller amounts of data. We used Excel to arrange the six sets of raw data from the surveys ensuring that the ten experimental conditions were presented in numerical order. We also used Google Sheets, which offers almost as many features as Microsoft Excel does, but has the advantage of making use of cloud computing, thus calculations are non-locally computed whereas Excel usually relies on the local hardware set-up.

The inferential statistics are done in form of a series of Chi-square tests, which in short can be described as a goodness-of-fit test: The Chi-square is a statistical test designed to identify the probability that the difference between the observed results, in our collected data, and the results that we were expecting to see, are due to chance or not. This means that we use the test measure the probability or level of significance for the rejection or acceptance of our H_0, and for this we set an alpha value of 0.05 or 5%. When rejection or accepting a H_0, we can simultaneously deduce that another hypothesis is true. The H_A is "*a statement indicating our belief in the nature and level of possible difference*" (Baines et al., 2002, p. 170), however rejecting the H_0 does not mean that the set H_A will directly be accepted.

Further, the upcoming data analysis and reporting of results will be done via descriptive statistics: These will be in the form of bar chart visualizations, which are used to create a more accessible and comprehensive summary for the discussion of the main features in the sample data.

4. Data Analysis

This chapter presents an analysis of the data gathered through the experimental research, presented in Chapter 2. The purpose of the analysis is to assess the PS and RQ's that was stated in the introductory chapter. Firstly, the nominal or raw data were sorted using Excel and Google Sheets, which were also used to produce the bar chart visuals presented later in this analysis. From here we went on to construct a statistical hypothesis H_0, a statistical way of formulating a prediction of our expected outcome (Study.com, 2015).

As mentioned in Section 3.6, we chose to analyse our data sets using a series of Chi-square test. In order to create a Chi-square test, we need to establish the following: a H_0, the observed results, the expected results, and the degrees of freedom. The H_0 is created previous to the test and the observed results are calculated in the aforementioned Excel sheets. To find the expected value we must multiply the row total instances with the column total and then divide it by the grand total. The degrees of freedom is found using $=(r - 1)(c - 1)$, where r is the total number of rows and c is the total number of columns (Baines et al., 2002). As we had to calculate multiple data sets, we do the Chi-square tests using an online form, which allowed an easy and uncomplicated calculation process.

4.1 Reporting Results

In order to examine the data from the survey design and thereby test whether or not the experimental conditions examined had any influence on the respondents intentions, a pair of H_0 and H_A is constructed. The H_0 will to be tested using a Chi-square test, and only in the cases where we cannot accept this hypothesis we will further discuss and apply the H_A. However, the H_A will appear along with the H_0 as one cannot stand without the other. All of the hypothesis pairs are linked to one specific RQ, however the RQ can be connected to more than just one pair. As we wanted to create a more accessible overview, the Likert scale labels used in the survey design were adapted to numeric values, which will be used in the Chi-square tests in Subsection 4.1.1 - 4.1.3. The numeric values are as follows: *Definitely* (1), *Probably* (2), *Neutral* (3), *Probably Not* (4), *Definitely Not* (5). Therefore the lower values will also be considered more positive in this analysis. In the bar chart visualizations, the Likert scale labels form have however been kept in their original form.

There are three values shown in each cell in the Chi-square test tables: The observed value collected in the study, the expected value which is displayed in Italic, and finally the frequency of the collected value shown in percentage. The latter are the values used to create the bar chart visualization, constructed for the purpose of providing a visual overview of the data results, and there will further be provided a summary of the comments made by respondents As the analysis entails numerous different variables and conditions being analysed we included these in the end of each subsection.

Finally, we have chosen to work with only two of the four possible points of inquiry made when examining each of the ten experimental conditions. The reasoning behind this choice is simply that a full report and analysis of all four points would create an overwhelming and very extensive chapter, with the risk of not conducting a proper in-depth analysis of each point. We therefore chose to look into the respondents' intention to purchase and recommend, as we find these two to be the most significant in the examination PS and RQs.

4.1.1 Type of Influence

Our second RQ was aimed at examining how and to what extend different types of social influence affects the consumers' purchasing decisions. From this we can posit the question of whether there is a relationship between the influence type and intent to purchase. As these were also included in our

study, we are operating with three different types of influence: *Ratings*, *Reviews*, and *Recommendations*. Our hypotheses are as follows:

H_0: There is no significant difference between the type of influence and the intent to purchase.
H_A: There is a relationship between the type of influence and the intent to purchase.

	Intent to purchase based on type of influence					
	1	2	3	4	5	Subtotals
Ratings	12 11.1 6,2%	56 39.1 28,7%	27 28.0 13,8%	60 54.6 30,8%	40 62.5 20,5%	195
Reviews	20 22.5 5,1%	77 79.0 19,5%	52 56.5 13,2%	104 109.9 26,4%	141 126.3 35,8%	394
Recommendations	24 22.5 6,1%	64 78.9 16,2%	62 56.51 15,7%	110 109.9 27,9%	134 126.3 34%	394
Subtotals	56	197	141	274	315	983

Table 4: Intent to purchase based on type of influence shown in Chi-square test

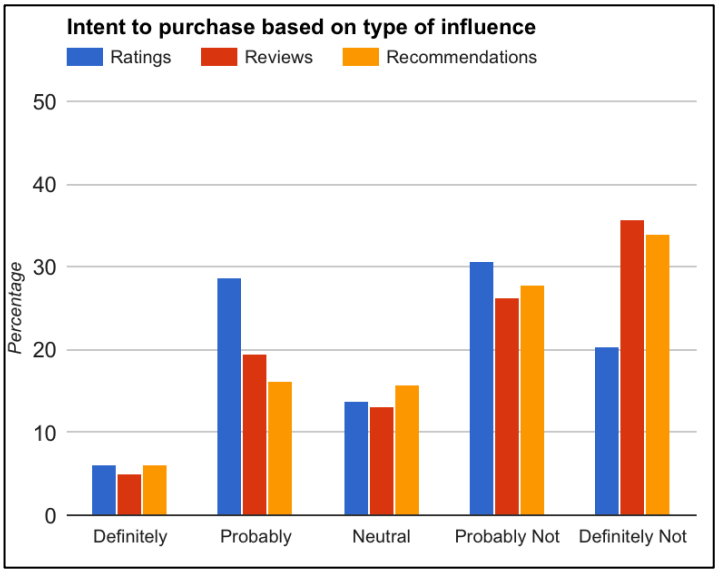


Figure 8: Intent to purchase based on type of influence as bar chart visualization

Using the values from Table 4, we created Figure 8, which shows the percentage of respondents' intent to purchase versus type of influence. We further calculated the median of *Ratings* as being 3,30, *Reviews* 3,68 and *Recommendations* 3,67, which indicate that the respondents show a slightly higher intent to purchase when being exposed to *Ratings* as a type of influence. However, the differences of 0,38 and 0,37 to the other two types of influence are so minor that it would be a stretch to conclude that *Ratings* are indeed more efficient. What is more interesting is the response pattern seen in the *Ratings* column (Table 4). Here the percentage of respondents who *Probably* want to purchase is

visibly higher compared to the other influences types. In the same way we can see that the *Ratings* column have more than 14% less of respondents stating that they *Definitely Not* have an intent to purchase, compared to *Recommendations* and *Reviews.* Looking to the H_0 we must however conclude that there is no statistically significant difference between the different types of social influence in terms of the consumer's intent to purchase as (χ^2 (8, N = 983) = 22.766, $p < 0.05$), and we can therefore accept H_0.

In order to extent the examination of RQ2 we further chose to investigate how different types of social influence can affect the consumer's decisions to recommend a product. The H_0 for examining this will be very similar to the aforementioned, as it will propose the question of whether there is a relationship between the influence type and this time, intent to recommend. We are again operating with the three types of influence: *Ratings*, *Reviews*, and *Recommendations*. The hypotheses here are:

H_0: There is no significant difference between the type of influence and the intent to recommend.
H_A: There is a relationship between the type of influence and the intent to recommend.

	Intent to recommend based on type of influence					
	1	2	3	4	5	Subtotals
Ratings	13 13.1 6,9%	32 24.1 17%	29 29.8 15,4%	68 49.6 36,2%	46 71.5 24,5%	188
Reviews	25 26.9 6,5%	53 49.3 13,8%	63 61.0 16,4%	90 101.5 23,4%	154 146.4 40%	385
Recommendations	29 27.0 7,5%	38 49.6 9,8%	60 61.3 15,5%	95 102.0 24,5%	165 147.1 42,6%	387
Subtotals	67	123	152	253	365	960

Table 5: Intent to recommend based on type of influence shown in Chi-square test

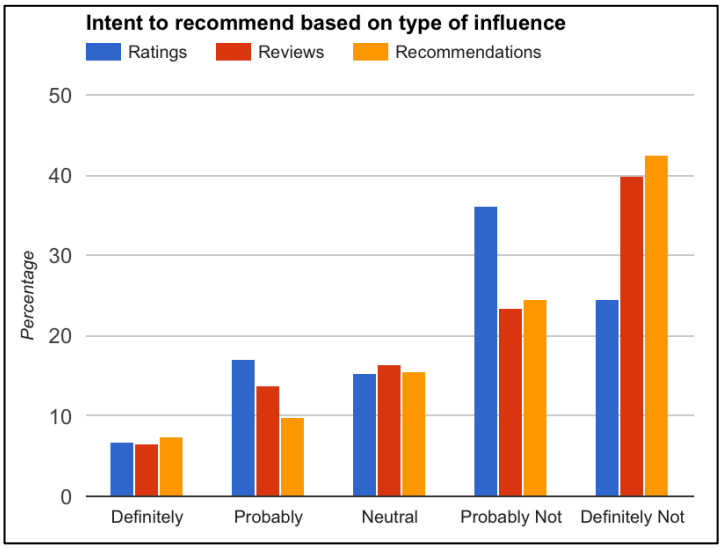


Figure 9: Intent to recommend based on type of influence as bar chart visualization

With the use of the data in Table 5 we were able to examine the respondents' intent to recommend a product based on the types of influence. This intent level was calculated to the median values of 3,54 for *Ratings*, 3,76 for *Reviews*, and 3.85 for *Recommendations*. We once again have to consider that a smaller value in this case equals an elevated level of intention to recommend. Just as in the first Chi-square test (Table 4) these median values give the indication that the respondents have more intent to recommend when *Ratings* are the type of influence. The differences to the other two median values are however even smaller than in the first test with a deviation of only 0,22 and 0,31, so here we felt the need to look to Figure 9, to make a closer comparison.

In the percentages shown in Figure 9 we do see that the respondents who answered that they would *Probably* want to recommend is slightly higher than the two other influences types. However, the difference is minor and the real deviation is in the amount of people whom would *Probably Not* and *Definitely Not* recommend the product. In the first instance of *Probably Not* intending to recommend, *Ratings* gave the strongest reaction. On the other hand a large percentage of respondents have no intention of recommending the product with *Reviews* or *Recommendations* as type. When we finally assess our H_0 we have to conclude that there again is no statistically significant difference between the effect of the three *Types of social influence* and the intent to recommend a product, as (χ^2 (8, N = 960) = 26.266, $p < 0.05$), and we therefore must accept the H_0.

After having reviewed the statistical and numerical data, we further looked into the comments posted by the main survey respondents (Appendix 5). For each comment posted we will refer to the type of influence used in the experimental condition where the comment was posted. We found that the majority of these had been provided when respondents had answered that they had little to no intention (*Probably Not* or *Definitely Not*) to purchase, recommend, share or save the product they had just reviewed. The in relation to comments, these were based on the things like the product's appearance, where one respondent stated that the product was not to his liking (*Recommendations*), to the lack of information presented e.g.: "*No actual information. What ports does it have? How many consoles can I plug in at once?*" (*Reviews*), and one respondent had concerns about delivery (*Ratings*). These comments suggest that it is not necessarily *Type of influence*, which is the most important factor to this sample population.

4.1.2 Source of Influence

Throughout this subsection of the analysis, we will be examining the relationship between the influence source and intent to purchase or recommend. This relates to our third RQ, which seeks to inquire how the source of social influence affects the customer's purchase decision. The statistical analysis in this subsection therefore focuses on the two independent variables of source *Customers* and *Friends*. The impact of these on the intent to purchase as well as to recommend is examined in the following paragraphs, specifically through the use of the Chi-square test displayed in Table 6 and Table 7. Alongside both of these tests we have further constructed two bar charts (Figure 10 and 11) that provide a visual overview of the observed results from our study. In the following paragraphs, the examination of the two independent variables for *Source of influence* and their respective levels of intent will enable us to examine the following hypotheses:

H_0: There is no significant difference between the source of influence and the intent to purchase.
H_A: There is a relationship between the source of influence and the intent to purchase.

	Intent to purchase based on source of influence					
	1	2	3	4	5	Subtotals
Customers	32 33.5 5,4%	130 117.8 22,1%	102 84.3 17,3%	155 163.9 26,4%	169 188.4 28,7%	588
Friends	24 22.5 6,1%	67 79.2 17%	39 56.7 9,9%	119 110.1 30,1%	146 126.6 37%	395
Subtotals	56	197	141	274	315	983

Table 6: Intent to purchase based on source of influence shown in Chi-square test

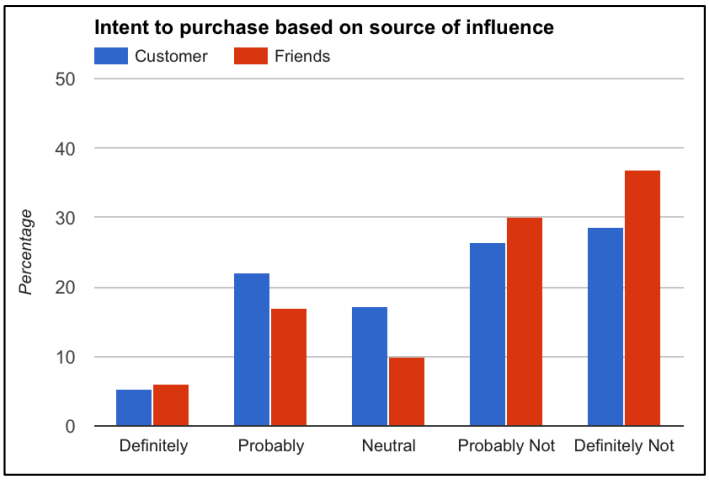


Figure 10: Intent to purchase based on source of influence as bar chart visualization

The numerical data generated in Table 6, will provide insights into the two variables effect on the respondents' intention to purchase. Looking to the median values of *Customers* and *Friends* in regards to their intent level shows the difference between these two is 0,60 points. For *Customers* the median is 3,51, whereas the value for *Friends* is 2,91. This could lead to the belief that the respondents were reacting more positively to the opinion information with *Friends* as the source, and the data in the bar

chart visualization (Figure 10) also point to the same assumption. Here, the percentages give an indication that when it comes to the intention to purchase respondents trust the opinion of *Friends* over *Customers*. The biggest leap can be seen in the columns with *Neutral* and *Definitely Not*. Here, the differences is respectively 7,4% and 8,3%, where the first instance shows the *Customers* to be the most influential, whereas the second suggest that *Friends* has a very negative influence on purchase intent.

In regards to the statistical evidence produced in the Chi-square test (Table 6) the percentages and median values do however not plays a significant role. In the assessment of the set H_0 we must deduce that there is no statistically significant difference between the effect of the two different *Sources of influence* and the intent to purchase as (χ^2 (4, N = 983) = 18.675, $p < 0.05$), and that we can thus accept our H_0.

As described in the beginning of this subsection, we also chose to look at the same *Sources of influence* and their respective effect on the respondents' intent to recommend. For the investigation of this another pair of hypotheses were set, where we posted the hypothesis that:

H_0: There is no significant difference between the source of influence and the intent to recommend
H_A: There is a relationship between the source of influence and the intent to recommend

	Intent to recommend based on source of influence					Subtotals
	1	2	3	4	5	
Customers	37 40.4 6,4%	85 74.2 14,7%	103 91.7 17,8%	152 152.6 26,3%	202 220.1 34,9%	579
Friends	30 26.6 7,9%	38 48.9 10%	49 60.3 12,9%	101 100.4 26,5%	163 144.9 42,8%	381
Subtotals	67	123	152	253	365	960

Table 7: Intent to recommend based on source of influence shown in Chi-square test

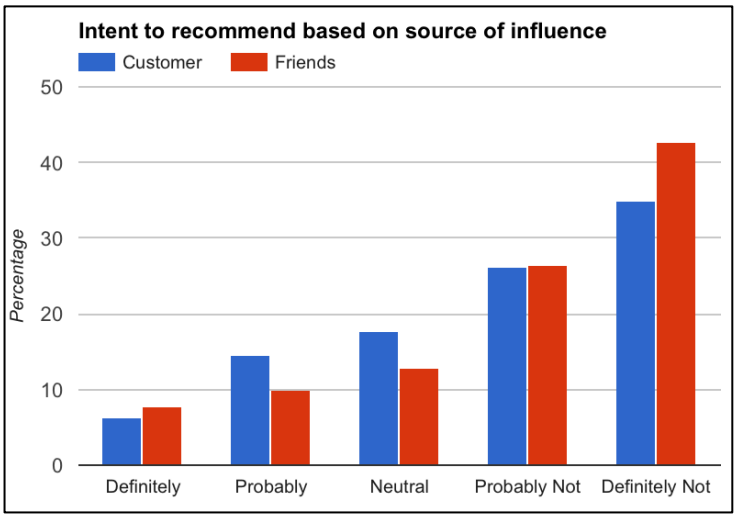


Figure 11: Intent to recommend based on source of influence as bar chart visualization

As our hypothesis describes, the data in this Chi-square test (Table 7) was gathered to examine if there was a variation between the intention to recommend between the two sources; *Customers* and *Friends*. This at first glance looks similar to the one concerning intention to purchase based on these two sources as influencers, though this is not the case: When we look at the median value for *Customers*, 3,68, it is not more than a 0,18-point difference to *Friends* with 3,86. These values are so close that if we based our analysis solely on the median value it would suggest that the two sources has an equal effect on the respondent. However, as we also examine other factors we can find that there is a minor difference between the two variables. When being exposed to an experimental conditions containing opinion information by *Customers*, we see that the respondents react more positive towards recommending the product. Nevertheless, in the negative end of the scale, the difference between *Customers* and *Friends* was still only 7,9% in *Definitely Not* and in *Probably Not* there is a mere 0,2% difference (Table 7 + Figure 11).

On the basis of the data that was calculated using the Chi-square test (Table 7), the evidence is somewhat similar to the findings made with use of the medians etc. When assessing H_0 we found that there was no statistically significant difference between the effect of the two different *Sources of influence* and the respondents intention to recommend a product: ($\chi^2(4, N = 960) = 11.995, p < 0.05$). Hereby we can accept our H_0.

Following the examination of the numerical and statistical data, we once more looked at the comments posted under both experimental conditions where *Customers* and *Friends* were the sources of influence. Here it was clear, just as in described in Subsection 4.1.1, that the opinion information provided by both *Friends* or *Customers* did seemingly not have a significant impact on the respondents. Instead their focus was on other elements presented on the mock-ups. In this instance some respondents reacted to *Friends* being the source with scepticism and irony "*Friends buying a piece of clothing would not influence me; I would rather trust a greater group of anonymous strangers*." (Appendix 5). When *Consumers* was the source no one reacted to this specifically, however many other reasons were given for not intending to purchase or recommend.

4.1.3 Interaction Between Source and Type

In this third subsection we investigate the last of our four RQs. This RQ inquires on the how the two independent variables, *Source of influence* and *Type of influence* interact and thereby affect the purchasing behaviour. For these analysis we coupled the independent variables in various ways, e.g. intention to purchase based on *Type of influence* with *Review* as source. This will then enable us to explore how, and whether the two variables would influence one another. In this subsection we will therefore be constructing eight different Chi-square tests to examine these interactions.

4.1.3.1 Influence type with Customers as Source

The initial two tests will look at the relation between intent to purchase or recommend based on influence type, having the *Customer* as the *Source of influence.* For this we post the following hypotheses:

H_0: There is no significant difference between the types of influence and the intent to purchase, when customers are the source of influence.
H_A: There is a relationship between the types of influence and the intent to purchase, when customers are the source of influence.

	Intent to purchase based on influence type and customers as source					
	1	2	3	4	5	Subtotals
Ratings	12 9.1 6,2%	56 42.0 28,7%	27 30.9 13,8%	60 54.0 30,8%	40 59.1 20,5%	195
Reviews	10 9.1 5,1%	40 42.2 20,4%	37 31.0 18,9%	50 54.3 25,5%	59 59.4 30,1%	196
Recommendations	5 8.8 2,6%	29 40.9 15,3%	28 30.1 14,7%	51 52.7 26,8%	77 57.6 40,5%	190
Subtotals	27	125	92	161	176	581

Table 8: Intent to purchase based on influence type and customers as source shown in Chi-square test

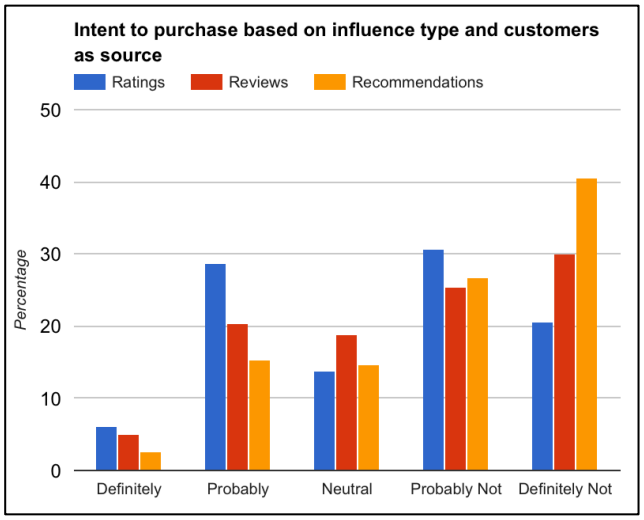


Figure 12: Intent to purchase based on influence type and customers as source as bar chart visualization

After computing the Chi-square test, we calculated the median values showing level of intention for each of the three *Types of influence*. Here we found that the level for *Ratings* was being 3.31, *Reviews* 3.55, and *Recommendations* 3.78, which again indicated that the respondents had a slightly higher intent to purchase when *Ratings* was the source. However, the median levels show that the differences are minor and all just above *Neutral*. With the use of the numerical data on percentages displayed in Table 8, we further constructed Figure 12, which provides an overview respondents' intent to purchase based on *Customers* as influence source. When closely examining the percentages displayed here, we do find that the respondents' intention to *Definitely* or *Probably* wanting to purchase the product is the highest in the *Ratings* column (Table 12). Compared to *Recommendations* alone, *Ratings* provoke an intention to purchase that is respectively 3,6% (*Definitely*) and 13,4% (*Probably*) higher. At the other end of the scale *Recommendations* cause 40,5% respondents to the *Definitely Not* intent to purchase, against only 30,1% for *Reviews* and 20,5% or *Ratings*. We once again look at our

H_0 for the Chi-square test, and have to must conclude that there is no statistically significant difference to be noted, between the types and the level of intent to purchase when *Customers* is the source, as (χ^2 (8, N = 581) = 26.525, p < 0.05). We therefore can accept the H_0.

From here we now wanted investigate the intention to recommend based on the same parameters as in the previous example. This means that we again have the three types of influence *Ratings*, *Reviews*, and *Recommendations* and the *Customers* as the source. From this we post the following statements in the hypotheses:

H_0: There is no significant difference between the types of influence and the intent to recommend, when customers are the source of influence
H_A: There is a relationship between the types of influence and the intent to recommend, when customers are the source of influence

	Intent to recommend based on influence type and customers as source					
	1	2	3	4	5	Subtotals
Ratings	13 12.0 6,9%	32 27.6 17%	29 33.4 15,4%	68 49.4 36,2%	46 65.6 24,5%	188
Reviews	12 12.4 6,2%	31 28.5 16%	44 34.5 22,7%	41 50.9 21,1%	66 67.7 34%	194
Recommendations	12 12.6 6,1%	22 28.9 11,2%	30 35.0 15,2%	43 51.7 21,8%	90 68.7 45,7%	197
Subtotals	37	85	103	152	202	579

Table 9: Intent to recommend based on influence type and customers as source shown in Chi-square test

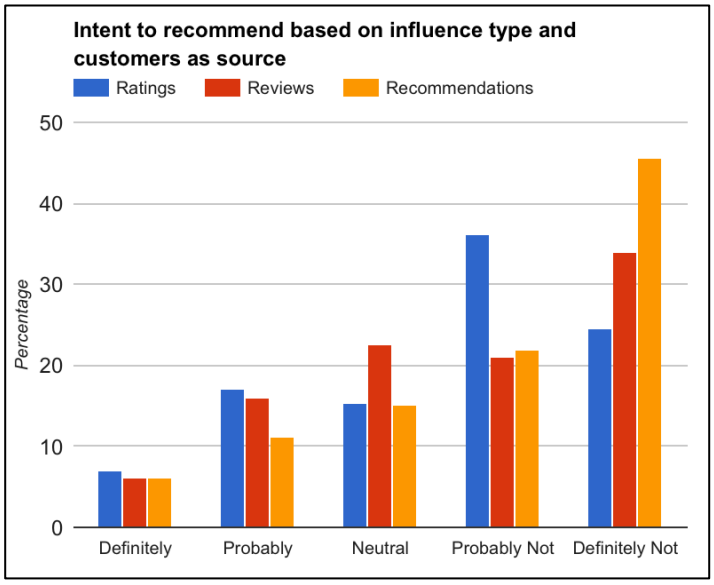


Figure 13: Intent to recommend based on influence type and customers as source as bar chart visualization

After creating the Chi-square test (Table 9), we initially calculated the median values for the three *Types of influence.* Keeping in mind that the lower values in Table 9 are a higher intention to recommend, whereas the higher values indicate a lower intention. The calculation of the medians showed us that *Recommendations* with a value of 3.88, gave the respondents an intention to recommend, which was in the low end of *Neutral.* Also neither *Ratings,* with a value of 3.54, nor *Reviews*, with the value 3.60, gave the respondents any higher intention to recommend. These minor differences between the three indicate that no matter which *Type of influence* the respondents were exposed to, they would never recommend the product. Further, Figure 13 shows the percentages of respondents intending to recommend a product, and here the clearest distinction is found in the negative end of the scale. Here respondents have no intention to recommend when seeing *Customer Recommendations* (45,7%), and would *Probably Not* recommend after seeing *Customer Ratings* (36,2%). Returning to Table 9 we look at the expected and observed values that is the basis for our statistical analysis. When looking at these we can say that there is no statistical significant difference between intention to recommend, whether *Ratings*, *Reviews* or *Recommendations* is the *type of influence* combined with *Customers* as the source: (χ^2 (8, N = 579) = 29.553, $p < 0.05$). Hereby we once more have accepted our H_0.

The comments for the experimental conditions where *Customers* were the source of influence (Appendix 5) ranged from lack of information (*Customer + Ratings*), return-costs being too expensive (*Customer + Recommendations*) and that the product was not adequate in regards to their needs: *"The specs of the TV. 120hz is insufficient" (Customer + Reviews).* From this it could be deducted that none of the three types, combined with *Customers* as source, had any real significant influence on the respondents.

4.1.3.2 Influence type with Friends as Source

Next we wanted to examine the interaction and relationship between the influence types now with *Friends* as the source. We must note that the variable *Friends* was only combined with two of the type variables; *Reviews* and *Recommendations.* Again we were to look investigate the respondents' intention to both purchase and to recommend the product in question, based on these influence types in combination with the variable *Friends*. We therefore posted this hypothesis pair for examination of purchase intention:

H_0: There is no significant difference between the types of influence and the intent to purchase, when friends are the source of influence
H_A: There is a relationship between the types of influence and the intent to purchase, when friends are the source of influence.

	Intent to purchase based on influence type and friends as source					
	1	2	3	4	5	Subtotals
Reviews	10 12.0 5,1%	37 33.6 18,7%	15 19.6 7,6%	54 59.7 27,3%	82 73.2 41,4%	198
Recommendations	14 12.0 7,1%	30 33.4 15,2%	24 19.5 12,2%	65 59.4 33%	64 72.8 32,5%	197
Subtotals	24	67	39	119	146	395

Table 10: Intent to purchase based on influence type and friends as source shown in Chi-square test

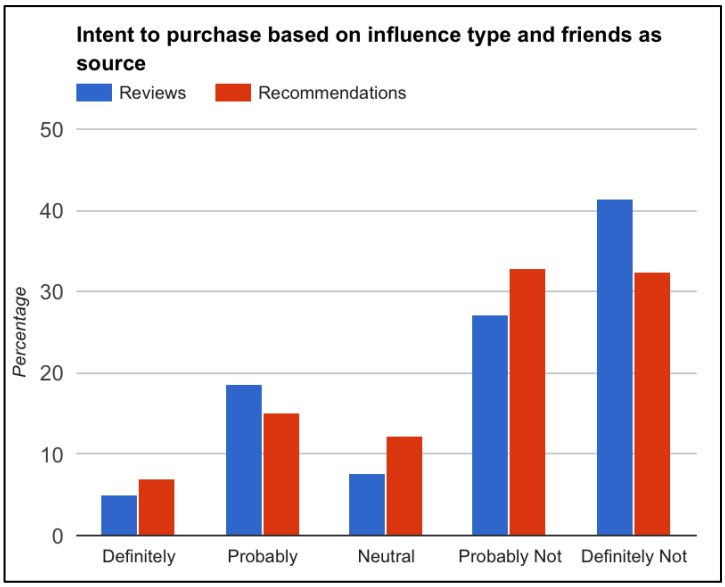


Figure 14: Intent to purchase based on influence type and friends as source as bar chart visualization

As this Chi-square test (Table 10) only contained two variables based on the *Type of influence*, we did not expect to see great differences in the results. When looking at the two median values there is not any large variation to detect. The median for *Recommendations* is at 3.64, very close to the *Reviews*, which has a value of 3.81. When knowing that a lower value here is considered as a stronger intention to purchase we can however see that *Recommendations* prompt a slightly more positive reaction from the respondents. When we further examine the percentages used to construct the bar chart visualization in Figure 14, it is even more evident that the two *Types of influence* cause an almost similar reaction pattern. The biggest difference is found in the columns showing an intention to *Definitely Not Purchase* the product in question. Here *Reviews* have a total of 41,4%, where *Recommendations* is at 33,5%. These numbers could indicate that in this particular case the median values give a rather accurate picture of the respondents' intention level.

The data described above and displayed in Figure 14, show us that there could be a slightly higher intention to purchase when exposed to *Reviews* from *Friends,* as opposed to *Recommendations* from the same source. However, looking to the statistical data from the Chi-square test we must again summarize that there is no statistical significant difference between the *Type of influence* when paired

with the source *Friends* (χ^2 (4, N = 395) = 6.708, $p < 0.05$), and that we therefore can accept our H_0. Now for the examination of the respondents intention to recommend, based on the same variables as described above in the example above, a similar hypotheses set was constructed:

H_0: There is no significant difference between the types of influence and the intent to recommend, when friends are the source of influence
H_A: There is a relationship between the types of influence and the intent to recommend, when friends are the source of influence

	Intent to recommend based on influence type and friends as source					
	1	2	3	4	5	Subtotals
Reviews	13 15.0 6,8%	22 19.1 11,5%	19 24.6 9,9%	49 50.6 25,7%	88 81.7 46,1%	191
Recommendations	17 15.0 8,9%	16 19.0 8,4%	30 24.4 15,8%	52 50.4 27,4%	75 81.3 39,5%	190
Subtotals	30	38	49	101	163	381

Table 11: Intent to recommend based on influence type and friends as source shown in Chi-square test

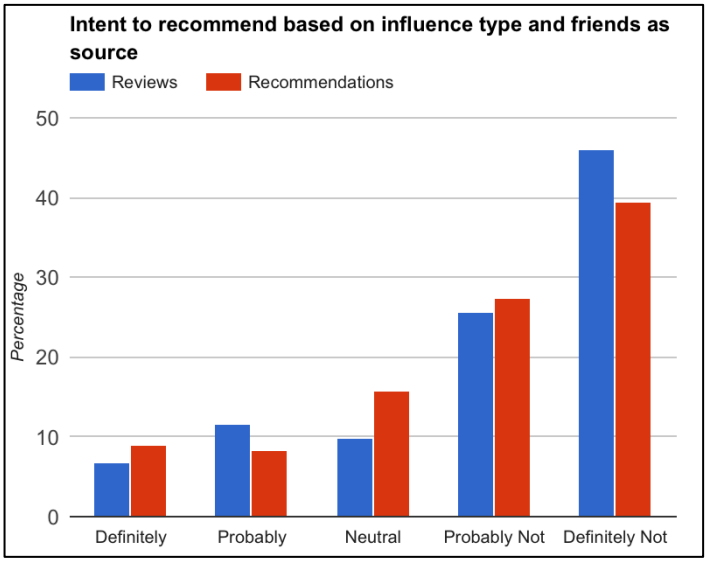


Figure 15: Intent to recommend based on influence type and friends as source as bar chart visualization

Here we started by examining the median values, as these were once again good indicators of how the respondents had reacted to the two different *Types of influence*. The median for *Reviews* was the highest of the two, and thereby the type, which caused the most negative reaction from the respondents. Here an entire 46,1% also stated that they would *Definitely Not* recommend this product in question (Table 11). The *Recommendations* also had a high median value of 3.76, and respectively 27,4% and 39,5% further responded that they would *Probably Not* or *Definitely Not*, recommend the

product here. This lack in intention to recommend becomes even more evident when the percentages are organized in the bar chart visualization in Figure 15, where we clearly see the negative tendency displayed. On the basis on the numerical data that was computed in the Chi-square, we further find that the difference between *Reviews* and *Recommendations* affect on the respondents' intention to recommend a product having *Friends* as source, statistically significant as (χ^2 (4, N = 381) = 5.073, p < 0.05). Here, we must once again accept our H_0.

We further looked at the comments from respondents when they were exposed to respectively *Reviews* and *Recommendations* (Appendix 5). These referred to everything from layout to description (*Friends + Recommendations*), as well as the lack of information (*Friends + Reviews*). What was especially interesting was the comment of one male respondent who said that he would not buy because the information on the mock-up had stated that eleven of his friends had also purchased the product (*Friends + Recommendations*). This statement indicates that *Friends* as the source of influence is not necessarily viewed as a positive factor when wanting to purchase a product.

4.1.3.3 Source of influence with Recommendations as Type

Lastly in this chapter, we would like to examine the relationship of the between influence source and the intent to purchase and recommend. As it was just the *Recommendations* and *Reviews* that were linked to the two sources *Friends* and *Customers*, we will only examine the effect on these two and not on the *Ratings.* First, we investigate the relation between the sources of influence and the intent to purchase when the influence type is recommendations, using the following hypotheses:

H_0: There is no significant difference between the sources of influence and the intent to purchase, when recommendations are the type of influence
H_A: There is a relationship between the sources of influence and the intent to purchase, when recommendations are the type of influence

	Intent to purchase based on influence source and recommendations as type					
	1	2	3	4	5	Subtotals
Customers	5 9.3 2,6%	29 29.0 15,3%	28 25.5 14,7%	51 57.0 26,8%	77 69.2 40,5%	190
Friends	14 9.7 7,1%	30 30.0 15,2%	24 26.5 12,2%	65 59.1 33%	64 71.8 32,5%	197
Subtotals	19	59	52	116	141	387

Table 12: Intent to purchase based on influence source and recommendations as type shown in Chi-square test

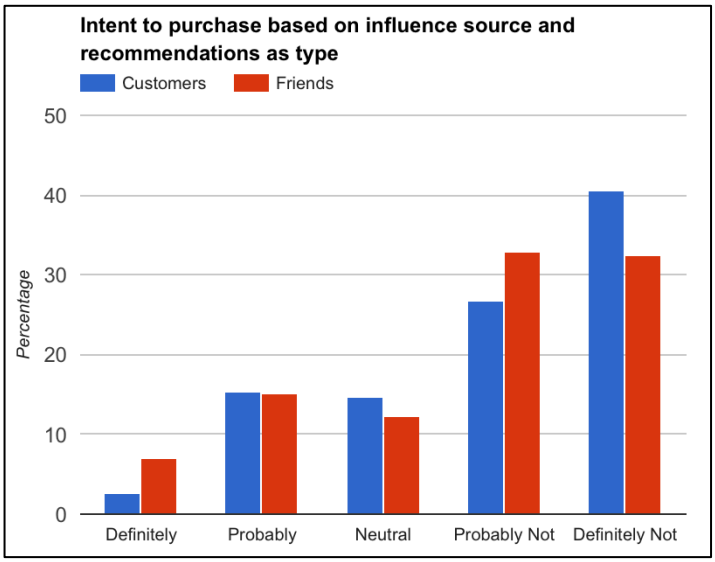


Figure 16: Intent to purchase based on influence source and recommendations as type as bar chart visualization

With the use of the numerical data in Table 12 we were able to examine the respondents' intent to purchase a product based on the source of influence, when *Recommendations* was the influence type. We started by looking at the median values for both sources and found that *Customers* had a value of 3.78, whereas *Friends* were at 3.65. We once more have to consider that the smaller value in this case equals a higher level of intention to purchase. So these median values indicate that there is a small, but almost insignificant difference between the two sources. Here we see that when *Friends* is the source of the *Recommendations* the intent is only 1,3 points higher than for *Customers*. The differences between the two variables are however very clear when looking to Figure 16.

Here we see that 7,1% of the respondents answered that they would *Definitely* purchase a product recommended by *Friends* whereas only 2,6% would purchase based on a recommendation by *Customers* in general. As the values at the other end of the scale show that the respondent were also less likely to purchase when viewing *Recommendations* from *Customers*, it would have been misleading to make any further assumptions or conclusion solely based on the percentages. When we finally assessed the stated H_0 we had to conclude that there is no statistically significant difference between the effect of the two *Sources of influence* and the intent to purchase a product, as ($\chi^2(4, N = 387) = 7.352$, $p < 0.05$), and we therefore can accept the H_0.

Next, we looked at the relationship between influence sources and intent to recommend when the influence type is recommendations. For this we constructed these hypotheses:

H_0: There is no significant difference between the sources of influence and the intent to recommend, when recommendations are the type of influence.
H_A: There is a relationship between the sources of influence and the intent to recommend, when recommendations are the type of influence.

	Intent to recommend based on influence source and recommendations as type					
	1	2	3	4	5	Subtotals
Customers	12 14.8 6,1%	22 19.3 11,2%	30 30.5 15,2%	43 48.4 21,8%	90 84.0 45,7%	197
Friends	17 14.2 8,9%	16 18.7 8,4%	30 29.5 15,8%	52 46.6 27,5%	75 81.0 39,5%	190
Subtotals	29	38	60	95	165	387

Table 13: Intent to recommend based on influence source and recommendations as type shown in Chi-square test

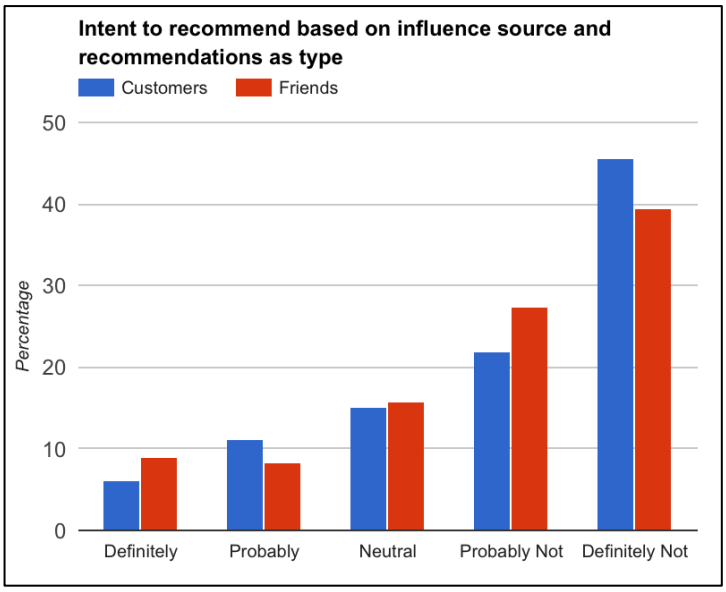


Figure 17: Intent to recommend based on influence source and recommendations as type as bar chart visualization

The respondents' intentions to recommend a product based on *Recommendations* from either of the two sources was in this case almost equal. The median for the intention level with the *Customers* as source is close to *Probably Not* wanting to recommend the displayed product, with a value of 3,89. For *Friends* the median was also close the same level of intention with the value 3,76. These value are in close proximity to information we were able view in Figure 17. Here it was clear to distinct that the majority of the respondents had none or little intention of recommending a product based on *Recommendations* from neither *Friends* nor *Customers*. For the latter, the percentage of respondents who would *Definitely Not* recommend was an entire 45,7%, giving *Friends* the advantage with 39,5%. If only looking at the percentages we could say that *Recommendations* from either source in general did not have a very positive effect on product recommendations. When assessing H_0 we found that there was no statistically significant difference between the two *Sources of influence* with recommendations as type and the effect they had on the respondents' intention to recommend as: (χ^2 (4, N = 387) = 3.900, $p < 0.05$). Hereby we could once more accept our H_0.

The respondents also provided comments to the experimental conditions where *Recommendations* was the influence type (Appendix 5) Two of the most conspicuous comments were both focused on the

type of opinion information. The first stated that there was: "*Not enough information to know if I would like it"* and continued by asking: "*...why did you hide the reviews?"*. The second simply stated that he would not purchase because there was no reviews (*Friends + Recommendations*). These two comments give us an indication that the *Type of influence* is for some consumers might be weighted more than who provides this the information.

4.1.3.4 Source of Influence with Review as Type

Our final two Chi-square tests will just as the preceding, be looking into the relationship between influence source and intent to purchase and recommend, now based on *Reviews* being the influence type. We set the following hypotheses to test the first instance, the effect on intention to purchase:

H_0: There is no significant difference between the sources of influence and the intent to purchase, when reviews are the type of influence.
H_A: There is a relationship between the sources of influence and the intent to purchase, when reviews are the type of influence.

	Intent to purchase based on influence source and reviews a type					
	1	2	3	4	5	Subtotals
Customers	10 9.9 5,1%	40 38.3 20,4%	37 25.9 18,9%	50 51.7 25,5%	59 70.1 30,1%	196
Friends	10 10.1 5,1%	37 38.7 18,7%	15 26.1 7,6%	54 52.3 27,3%	82 70.9 41,4%	198
Subtotals	20	77	52	104	141	394

Table 14: Intent to purchase based on influence source and reviews a type shown in Chi-square test

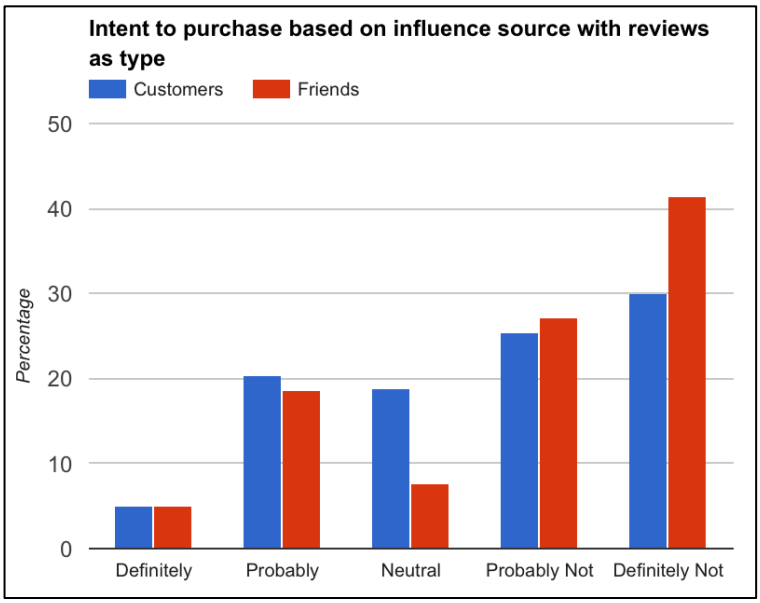


Figure 18: Intent to purchase based on influence source and reviews a type as bar chart visualization

The numerical data generated in Table 14, provides knowledge of the two variables effect on the respondents' intention to purchase, when *Reviews* is the influence type. Looking at the median values of *Customers* and *Friends* respectively a mere difference of 0,27, as the value for *Customers* 3,55 and for *Friends* 3,82. This could lead to the deduction that the respondents reacted more positively to the *Reviews* posted when *Customers* were the source. The values in Table 14, which have also been visualized in the bar chart (Figure 18), point to the same assumption. Here, the level of intention to purchase is significantly higher for *Customers* than it is for *Friends*. However, the *Sources of influence* are almost consecutive terms of percentages except for in *Neutral* and *Definitely Not.* In the first case the *Customers* effect the respondents the most with 18,9% against 7,6%, whereas the *Friends* influence is 11,3% higher in the second.

In regards to the statistical evidence produced in the Chi-square test (Table 14) the percentages and median values however does not play any statistically significant role. In the evaluation of our H_0 we must state that there is no statistically significant difference between the effect of the two different *Sources of influence* and the intent to purchase as (χ^2 (4, N = 394) = 13.320, $p < 0.05$), and we therefore accept our H_0.

As mentioned in the previous paragraphs, this twelfth and final Chi-square was computed to test the association between the influence source and respondents intention to recommend, having *Reviews* as the influence type. Our last hypotheses pair therefore presents the statement that:

H_0: There is no significant difference between the sources of influence and the intent to recommend, when reviews are the type of influence.
H_A: There is a relationship between the sources of influence and the intent to recommend, when reviews are the type of influence.

	Intent to recommend based on influence source and reviews as type					
	1	2	3	4	5	Subtotals
Customers	12 12.6 6,2%	31 26.7 16%	44 31.8 22,7%	41 45.4 21,1%	66 77.6 34%	194
Friends	13 12.4 6,8%	22 26.3 11,5%	19 31.3 9,9%	49 44.7 25,7%	88 76.4 46,1%	191
Subtotals	25	53	63	90	154	385

Table 15: Intent to recommend based on influence source and reviews as type shown in Chi-square test

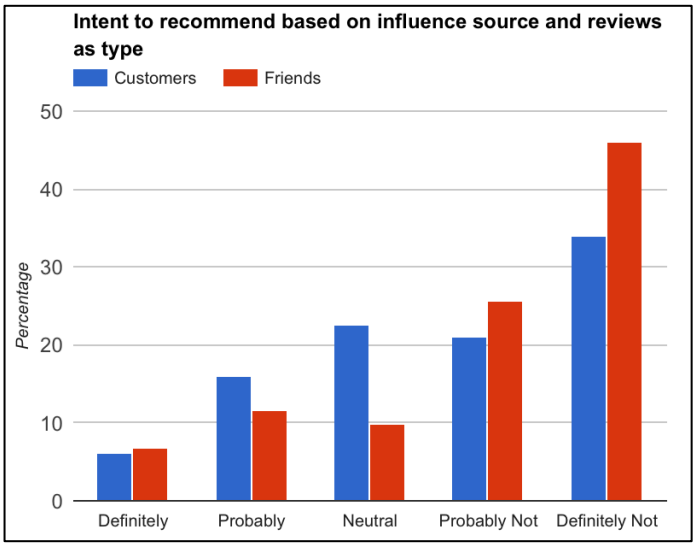


Figure 19: Intent to recommend based on influence source and reviews as type as bar chart visualization

As our hypothesis describes, the data in this Chi-square test (Table 15) was computed in order to examine if there was any variation between the respondents intention to recommend when *Reviews* came from the two sources *Customers* and *Friends.* The median value for *Customers* was 3,61 against the value for *Friends was at* 3,92. The percentages used to construct the bar chart (Figure 19), does nonetheless indicate some prominent differences between the two *Sources of influence.*

When being exposed to an experimental condition containing *Reviews* by *Customers*, we saw respondents reacting more positive towards recommending the product in question. Here 16% would *Probably* recommend, while only 11% of respondents exposed to *Friends* as source would do so. In the negative end of the parameter, the difference between percentages of respondents who would *Definitely Not* recommend is 12,1% *Friends* having most negative effect. On the basis of the data that was computed using the Chi-square test (Table 15) assessed our set H_0 and found that there was no statistically significant difference between in the effect of *Reviews* posted by *Customers* and *Friends* in the respondents intention to recommend a product: (χ^2 (4, N = 385) = 15.320, $p < 0.05$). Hereby we can also accept our final H_0.

In terms of comments for this section, the respondents had a very similar reaction pattern when exposed to *Reviews*. Both conditions with *Customers* and *Friends* as the source of influence, had prompted comments on the lack of information on the product such as ”*The abundance of information; I was overwhelmed”* (*Friends + Reviews*, Appendix 5). Further, respondents noted that they would never purchase or recommend a DVD or movie as they would “*...Not buy a film in dvd (sic!), but as digital download*” or “*...just bootleg it*” (*Customers + Reviews*, Appendix 5). This is a good example of the fact that some products have an expiration date or are being replaced by other technologies. Here, the physical version of the product, a DVD, have been replaced by a service allowing the consumer to stream or download the film onto their computer or tablet. Here it is possible that another in type of opinion information or source will not change the outcome.

Before entering the fifth and final chapter of this Master thesis, we must add a comment in regards to our just finished analysis: Our survey design contained four points to answer via the Likert-scale after each experimental condition had been displayed. The first two inquiries were about the respondents’ intention to purchase and recommend the product they had been shown. As stated in the beginning of this chapter, the results from these two points have been analysed and discussed, whereas the two other points *Share* and *Save*, have not. However, we did compute the Chi-square tests, as well as

construct the bar chart visualisations for these points, but in none of these 12 tests did we find there to be any significant differences to the ones posted in the analysis above. All of the results can be viewed in Appendix 6.

5. Discussion and Conclusion

This final chapter contains a discussion, as well as thoughts on what the experimental research conducted throughout this Master thesis has entailed. We will also discuss the limitations of this specific type of research, as well as have a brief look at the possible perspectives of future research into the phenomenon of social influence in e-Commerce.

In the introduction in Chapter 1, we stated that the main motivation for examining the phenomenon of social influence in e-Commerce was a: "*...Hope to create an interest in, and shed light on the effects of social influence, and thereby push further research in the area.*" (Section 1.4) As it is difficult to establish whether or not the first point, creating an interest in the phenomenon, has in fact been reached, we want to focus primarily on the second and third point. Beforehand it still must be noted, that we believe social influence to be, not as an effect of our specific research, but in general a phenomenon, which will continue to attract an increasing amount attention in the coming years. We do also believe that even though the extend of the experimental research done in this thesis is not wide, it has shed light on the effects of social influence, or maybe lack thereof.

Through the analysis and reporting (Chapter 4) of the results collected through our survey design (Section 3.2) we detected a rather distinct tendency in the commentary pattern: Many respondents provided comments which indicated that the opinion information they had been exposed to was insignificant to their choices. This supported the amount of negative responses given to the inquiry about intention to purchase and recommend the product in question. The insights provided through the analysis of the data further indicated that our experimental research design was possibly not the most adequate for testing this particular. This itself is a valuable insight, as it will be useful in possible further analysis on the same field, saving effort and time. This is also why we view the field of inquiry as being one where researchers, as well as companies, in the future would be able to gain great and valuable insights, into the effect of opinion information. However, in order to obtain this knowledge concerning the extent of the effect caused by the social influence of opinion information in e-Commerce, a more in-depth study would be necessary.

This future research of the phenomenon should however entail a wider array of data collection methods. Here we, as mentioned in the scope (Section 1.5), deem methods such as single-subject interviews, eye- or mouse tracking testing, and A/B or multi-variant testing, as possibly being suitable methods. Further implementing these tests and methods over a longer period of time would provide the researcher with a deeper understanding of the phenomenon of social influence. Ideally, this kind of testing would be done using an already existing e-Commerce shop with an extensive and stable amount of traffic to be evaluated, to finally be able to distinguish to a higher degree of precision the results in view. Future examination of the phenomenon could also be coupled with areas such as cognitive and psychological research. Further, an in-depth investigation of the other sources of influence, such opinion makers like magazine etc. influence buying behaviour and if their persuasive power is stronger than that of peers. As social structure and norms in terms of influence and purchasing behaviour may vary enormously, the outcome will incidentally further differ if a different group was tested.

Even though our results did not reveal any greater conclusions or insights into the phenomenon of social influence, we do see this thesis paper is making some contributions to the literature on social influence. The initial contribution lies in the investigation into the purchasing intention of consumers on e-Commerce websites, when being exposed to social influence in form of opinion information. Here we found that none of the different experimental conditions, and thereby the independent variables, stood out as being significantly more influential than others.

From our examination of the data we further learned, that when the respondents did respond to the opinion information, their reactions were very similar. A clear example was that the information stemming from friends prompted a much more negative reaction pattern, then when it came from customers. The reactions to the independent variable ratings, was correspondingly ranging in the negative end of the scale when compared to the other two types of influence, recommendations and reviews. The comments and inputs provided by respondents further showed us that many of the respondents did not find any exceptional value in the opinion information, and was more concerned with the price, value or brand of the product they were viewing. Although this information is not valid as statistical evidence, it does point to the fact that social influence might not be an important factor to many consumers, at least it did not seem to be something to which they give much attention or thought.

Literature list

Arala, S., Muchnika, L., Sundararajana, A., 2009, Distinguishing Influence-Based Contagion From Homophily-Driven *Diffusion In Dynamic Networks*, PNAS, vol. 106 no. 51, Washington, USA

Baines, P., Chansarkar, B., 2002 *Introducing Marketing Research*, John Wiley & Sons, New Jersey, USA

Barnum, C.M., 2011, *Usability Testing Essentials: Ready, Set... Test!*, Elsevier, Burlington, USA

Bearden. W.O., Etzel, M.J, 1982, *Reference Group Influence On Product And Brand Purchase Decisions*, Journal of Consumer Research vol. 9

Bias, R.G., Mayhew, D.J., 2005, *Cost-Justifying Usability: An Update for an Internet Age*, Elsevier, San Fransico, USA

Bond, R.M., Fariss, C.J., Jones, J.J., Kramer, A.D.I., Marlow, C., Settle, J.E., Fowler, J.H., 2012, *A 61-Million-Person Experiment In Social Influence And Political Mobilization*, Nature, USA

Bonfield, E.H., 1974, *Attitude, Social Influence, Personal Norm, and Intention Interactions as Related to Brand Purchase Behavior*, Journal of Marketing Research, Vol. 11, No. 4

Bredley, S., 2010, *Designing For A Hierarchy Of Needs*, Smashing Magazine, Germany

Brohan, M., 2014, *Nike's web sales grow 70% in Q1*, Internet Retailer

Caffey, D., Ellis-Chadwick, F., Johnston, K., Mayer, R., 2006, *Internet Marketing*, Pearson Education Limited, Essex, UK

Carli, L.L., 2001, *Gender and Social Influence*, Journal of Social Issues

Charness, G., Gneezy, U., Kuhn, M.A., 2011, *Experimental methods: Between-subject and within-subject design*, Journal of Economic Behaviour & Organization 81, Elsevier, USA

Childers, T., Rao, A., 1992, *The Influence Of Familial And Peer-Based Reference Groups On Consumer Decisions*, Journal of Consumer Research, Vol 19

Christiansen, B. 2014, *Handbook Of Research On Effective Marketing In Contemporary Globalism*, Business Science Reference, Hersey, USA

Cialdini, R.B, Goldstein, N.J., 2004, *Social Influence: Compliance and Conformity Annual Reviews Psychology*, USA

Cisco, 2014, *Cisco Visual Networking Index: Forecast and Methodology: 2013–2018*, Cisco, California, USA

Cisco, 2012, *Cisco Visual Networking Index Forecast Projects 18-Fold Growth in Global Mobile Internet Data Traffic From 2011 to 2016*, Cisco, California, USA

Creswell, J.W., 2007*, Qualitative Inquiry & Research Design: Choosing Among Five Approaches*, Second edition, Sage Publications, California, USA

Creswell, J.W., 2014, *Research Design: Qualitative, Quantitative, and Mixed Methods Approaches*, Fourth Edition, Sage Publications, California, USA

Dooley, R., 2012, *Brainfluence: 100 Ways to Persuade and Convince Consumers with Neuromarketing,* John Wiley & Sons, New Jersey

Dulany, D. E., 1968, *Awareness, Rules And Propositional Control: A Confrontation With S-R Behavior Theory*" in Dixon, T.R., Horton, D.L. (Eds.), *Verbal Behavior And General Behavior Theory*, p. 340-387., Prentice Hall, New Jersey, USA

e-marketer.com, 2014, *Social Networking Reaches Nearly One in Four Around The Worlds*, e-marketer.com http://www.emarketer.com/Article/Social-Networking-Reaches-Nearly-One-Four-Around-World/1009976

Evans, K., 2015, *H&M Makes A Big E-Commerce Push*, Internet Retailer

Farell, S., 2011, *Social Influence Benefits The Wisdom Of Individuals In The Crowd*, PNAS, no. 36, Washington, USA

FDIH, 2015, *FDIH e-handelsanalyse Årsrapport 2014*, FDIH, Copenhagen, Denmark

Fogg, B.J., 2009, *A Behavior Model For Persuasive Design*, Stanford Uni. Persuasive Technology Lab, Stanford, USA

Friedkin, N.E., Johnson, E. C., 1999, *Social Influence Networks And Opinion Change*, Advances in Group Processes

Germar, M., Schlemmer, A., Krug, K., Voss, A., Mojzisch, A., 2013, *Social Influence And Perceptual Decision Making: A Diffusion Model Analysis*, Personality and Social, Psychology Bulletin

Giddens, A., 1987, *Social Theory And Modern Sociology*, Stanford University Press, UK

Gregor, S., 2006 , *The Nature Of Theory In Information Systems*, MIS Quarterly

Hogg, T., Lerman, K., 2014, *Effects Of Social Influence In Peer Online Recommendation*, California, USA

Hundley, V., van Teijlingen, E., 2002, *The Importance Of Pilot Studies, In Nursing Standard*, p. 33-36, ProQuest Nursing & Allied Health Source vol. 16, no. 40, Aberdeen, Scotland

Hursh, S.R., 1984, *Behavioural Economics In Journal Of Experimental Analysis Of Behaviour*, p. 435-452, Journal of Experimental Analysis of Behaviour no. 3

Interactive Advertising Bureau, 2008, *User Generated Content, Social Media, And Advertising — An Overview Interactive Advertising Bureau*

Ioannides, Y. M. , 2012, *From Neighborhoods To Nations*, chapter 2, 3 & 9, Princeton University Press, USA

ITU, *Persuasive Design: Old Wine In New Bottles,* ITU, Copenhagen, Denmark

Kallehave, O., Skov, M.B., Tiainen, N., 2011, *Persuasion In-Situ: Shopping For Healthy Food In Supermarkets*, Department of Computer Science, Aalborg University, Aalborg, Denmark

Kelly, D., 2009, *Methods For Evaluating Interactive Information Retrieval Systems With Users*, Foundations and Trends in Information Retrieval Vol. 3, No. 1–2

Kim, Y.A., Srivastava, J., 2007, *Impact of Social Influence In E-Commerce Decision Making*, Minnesota, USA

Lah, N.C.L, Hussin, A.R.C, Dahlan, H.M., 2013, *Improving B2C e-Commerce Trust through Social Presence Factors*, in Yama, F. (Ed.) ICCRD, 5th, Journal Of Information Systems Research And Innovation, Ho Chi Minh City, Vietnam

Lazar, J., Feng, J.H., Hochheiser, H., 2010, *Research Methods In Human-Computer Interaction*, John Wiley & Sons, New Jersey, USA

Lewis, K., Gonzalez, M., Kaufmanb, J., 2012, *Social Selection And Peer Influence In An Online Social Network*, Proc Natl Acad Sci, USA

Li, S., 2013, *Your Friends Like It. Do You? Examining Social Recommendation System's Effects On Consumer Preference Similarity*, Siyuan Li, HongKong, China

Lockton, D., Harrison, D., Stanton, N., 2008, *Design with Intent: Persuasive Technology in a Wider Context*, in H. Oinas-Kukkonen et.al *Persuasive*, Springer, Berlin, Germany

Lorenz, J., Rauhut, H., Schweitzer, F., Helbing, D. 2011, *How Social Influence Can Undermine the Wisdom*, PNAS, no. 22, Washington, USA

Maynard, D.M., 1991, *Goffman, Garfinkel, and Games*, Sociological Theory, Vol. 9, No. 2

McFerran, B., Dahl, D.W., Fitzsimons, G.J., Morales, A.C., 2009, *I'll Have What She's Having: Effects Of Social Influence And Body Type On The Food Choices Of Others*, Journal of Consumer Research Vol. 36

McKay, L., Pearson, L., Peart, M., Utreras, M., Wang, H., 2011, *Gen X: Demographic Profile And Marketing Approaches*, e-Marketer.com

Meyrowitz, J., 2005, *The Rise Of Glocality New Senses Of Place And Identity In The Global Village* in Nyiri, K. (Ed.) *A Sense Of Place: The Global And The Local In Mobile Communication*, Passagen Verlag. Vienna, Austria

Muchnik, L., Aral, S., Taylor, S.J. 2013, *Social Influence Bias: A Randomized Experiment*, Science, Washington, USA

O'Brien, H. L., Toms, E. G., 2008, *What Is User Engagement? A Conceptual Framework For Defining User Engagement With Technology*, Journal of the American Society for Information Science and Technology, Nova Scotia, Canada

Page, R.E., Kray, C., 2010, *Ethics And Persuasive Technology: An Exploratory Study In The Context Of Healthy Living,* 1st int. Workshop on Nudge & Influence Through Mobile Devices, Lissabon, Portugal

Patton, M.Q., 1990, *Qualitative Evaluation And Research Methods*, p. 169 - 186 Sage Publications Beverly Hills, USA

Patton, M.Q., 1999, *Enhancing The Quality An Credibility Of Qualitative Analysis*, Health Services Research

Pickard, A.J., 2013, *Research Methods In Information*, Second Edition, Facet Publishing, London, UK

Qualman, E., 2009, *Socialnomics: How Social Media Transforms The Way We Live And Do Business*, John Wiley & Sons, New Jersey, USA

Redström, J., 2006, *Persuasive Design: Fringes And Foundations*, in Persuasive p. 112 – 122, Springer-Verlag, Berlin, Germany

Revilla, M.A., Saris, W.E., Krosnick, J.A., 2014, *Choosing The Number Of Categories In Agree–Disagree Scales,* p. 73-97, Sociological Methods Research vol. 43 no. 1, USA & Spain

Riva, G., Galimberti, C., 1998, *Computer-Mediated Communication: Identity And Social Interaction In An Electronic Environment In: Genetic, Social And General Psychology Monographs*, p. 434-464, Heldref Publications

Sandvik, K., 2009*, Researching Online Worlds: Challenging Media And Communication Studies*, Mediekultur vol. 25, no. 47, Denmark

Schuz, B., 2014, *Wheel of Persuasion*, Converrsion Jam #4, Stockholm, Sweden

Sherif, M., 1935, *A Study Of Some Social Factors In Perception*, Archives of Psychology, 27, 17-2

Sonne-Ragans, V., 2012, *Anvendt Videnskabsteori - Refleteret Teoribrug I Videnskabelige Opgaver*, Samfundslitteratur, Frederiksberg, Danmark

Statista, 2014, *B2C E-Commerce As Percentage Of Global GDP From 2009 To 2018*, Statista, USA

Statista, 2014, *Biggest Internet Companies In Selected Countries As Of June 2014, Based On Value (In Million U.S. Dollars)*, Statista, USA

Strauss, A. L., Corbin, J. M., 1998, *Basics of Qualitative Research: Techniques and Procedures for Developing Grounded Theory*, Sage Publications, California, USA

Tapscott, D., 2009, *Grown Up Digital - How The Net Generation Is Changing The World*, McGraw-Hill

The Economist 2012, *Nudge Nudge, Think Think: The Use Of Behavioural Economics In Public Policy Shows Promise*,

The Economist, USA
Tullis, T.S, Albert, W., 2008, Measuring the User Experience: Collecting, Analyzing, and Presenting Usability Metrics, Morgan Kaufman
Underhill, P., 2009, *Why We Buy - The Science Of Shopping: Updated And Revised For The Internet, The Global Consumer And Beyond*, Simon & Schuster Paperbacks, New York, USA
van Teijlingen, E., Hundley, V., 2002, *The Importance Of Pilot Studies*, Nursing Standard
Venkatesh, V., Morris, M.G., 2000, *Why Don't Men Ever Stoop To Ask For Directions? Gender, Social Influence, and Their Role in Technology Acceptance and Usage Behaviour*, MIS Quarterly
Wang, A., Zhang, M., Hann, I., 2014, *Socially Nudged: A Quasi-Experimental Study of Friends' Social Influence in Online Product Ratings*, Michael Zhang Blog
Whitaker, K., 2014, *Reducing Abandoned Shopping Carts In E-Commerce*, Smashing Magazine, Germany
Wood, L., Egger, M., Gluud, L.L., Schulz, K.F., Jüni, P., Altman, G.D., ... Sterne, J.A.C., 2008, *Empirical Evidence Of Bias In Treatment Effect Estimates In Controlled Trials With Different Interventions And Outcomes: Meta-Epidemiological Study*, The BMJ
Wren, Kevin., 2002, *Social Influence – Social Pschycology,* Routledge, London, UK

Websites:

http://amazon.com, 2015
http://facebook.com, 2015
http:// reddit.com/r/Assistance/, 2015
http://reddit.com/r/Favors/, 2015
http://reddit.com/r/SampleSize/, 2015
http://study.com, 2015
http://surveygizmo.com, 2015
http://zalando.com, 2015
http://turner.faculty.swau.edu/mathematics/math241/materials/contablecalc/, 2015
http://twitter.com, 2015

Appendix

Appendix 1: Company interview

We asked a long line of questions, but our contact at the company in question was only able to answer some of them. These are the ones included in this appendix. The interview questions where answered via email

General:
Interviewer: To what extend is user generated content used in e-commerce?
Company representative: *User Generated Content (UGC) is used on every corner of e-commerce. If you look at a modern web shop, you will find lots of interaction points for users - be it the omnipresent Like and Tweet boxes, simple comments on a product page, or even full reviews of expensive technical hardware.*

I: How valuable do you think data, and user data in specific, is for an e-commerce company?
CR: *User data is extremely valuable to e-Commerce company. It helps serving a user exactly what she or he wants. One big problem for web shops is always that they cannot see exactly who their customer is - if it is a 60-year-old man from a little place in North America or a 15 year old girl from the suburbs of Tokyo. If you have a regular shop on a street, you pretty will be able to observe who is coming to the shop and what they are looking for, also how long they stand in front of which article and turn it around and around until they do or do not buy it for some reason. A web shop needs to find this out by mining data.*

I: What role does social media play for Zalando's e-commerce business?
CR: *For Zalando, social media is a valuable brand building channel. With this Zalando is able to show their followers what they stand for, how the 'human face' of the online shop looks like. But in Zalando's shop you will also find comments and recommendations on sizes etc. on the product pages, as well as UGC on content pages (like videos or pictures). Also, we provide the option for users to directly interact with Zalando's support team via Facebook.*

Specific:
I: What different social media touch points do you use for e-commerce purposes?
CR: *Zalando uses most high-frequented social networks (e.g. Facebook, Instagram etc.), also direct interaction via recommendation functions are available, and of course also ratings of products. You can basically access a social network from every page on a Zalando website.*

I: What have you seen as benefits and disadvantages in terms of the different touch points?
CR: *Zalando knows that recommendations or comments on product pages can be helpful in terms of lowering a return rate (e.g. if a manufacturer produces larger or smaller than usual sizes in articles). It is also being tested whether comments and shares increase sales for articles, meaning if users are more likely to buy a product if it has high ratings or was recommended by users connected to them. Unfortunately, I cannot go into further detail about exact results and leverages.*

Social media:
I: Does the existence of user generated content influence online buying behavior? Influences purchases?
CR: *From what I've seen, there is not a general yes or no to this question. It is very much depending on the product group, the individual product, and of course of the sort of UGC. What seems to influence conversions are for example detailed user reviews on electronic/entertainment products, e.g. TVs, Monitors, Cameras, Computer Games, and Music. But also recommendations on sizes for clothing or shoes (e.g. too small in usual size, etc.) has an influence on conversions - and also on return rates. So, if I can read something like "very soft abric, but it's a lot too small", I know that I have to buy it in a bigger size. Good for me as a customer because I don't have to return it, good for me as a retailer because I don't have to pay twice for shipping and handling. Also recommendations on quality of an article.*

For a webshop, this information is highly valuable since it can be used for sorting products better or generating content which serves as a general guide. This in the end can be also communicated through the different social networks and thus provoking extra UGC. What can be seen at the moment is that some major web shops, amongst those Zalando, are currently testing if how much impact presenting user touch points and UGC actually has - and if it's worth aiding this by providing an incentive for users to generate content.
Something that is related to this topic but less visible is of course the influence of buying behaviour on the position of articles in the catalog. If an article is being viewed or bought frequently, it will be rated higher in the catalog and thus visible to more users.

I: If so - how?
CR: See above

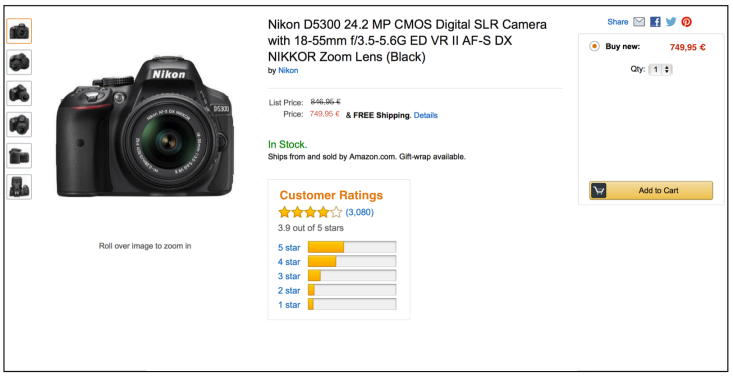
I: If not - why?
CR: See above

Future:
I: What would you like to do? User generated content etc.
CR: *There are always some ideas around, not all of them are actually paying off, of course. For instance, providing a tool for customers to create their own complete outfits (e.g. you can add a black silk shirt, a matching pair of white sneakers, some silver bracelets and a black bag, and a coat to the black leather skirt that you just liked) by only a few clicks that they can share with their friends or a community, and comment it further, maybe alter it a bit or create a matching version for a man - that would give users a way to become engaged with the web shop on a whole different level. Another idea might be to give users a good incentive to give qualified feedback about the article that they just bought. Or a third idea would be to give users the possibility.*

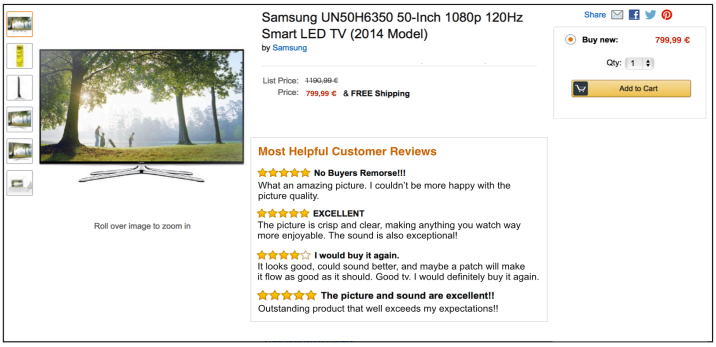
Appendix 2: Interface Mock-up Designs

1. *Ratings* by *customers* on a *low* priced product

2. *Reviews* by *customers* on a *low* priced product

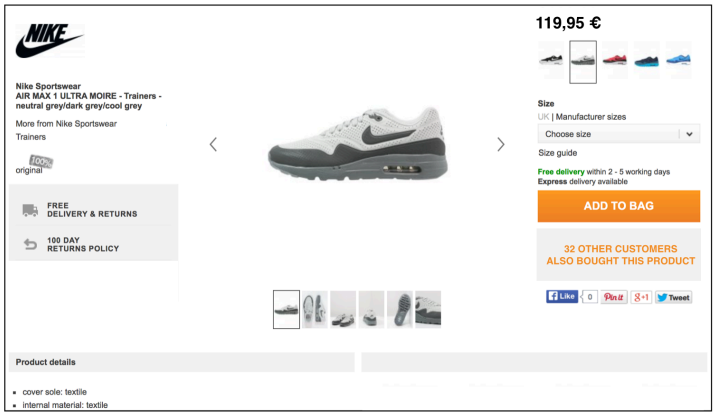


3. *Ratings* by *customers* on a *high* priced product

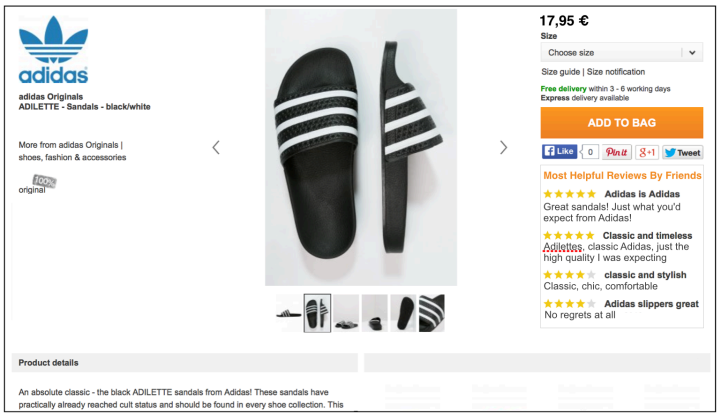


4. *Reviews* by *customers* on a *high* priced product

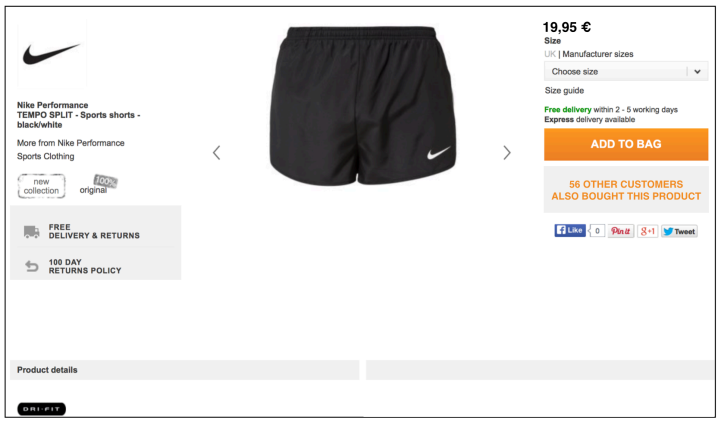
5. *Reviews* by *friends* on a *high* priced product



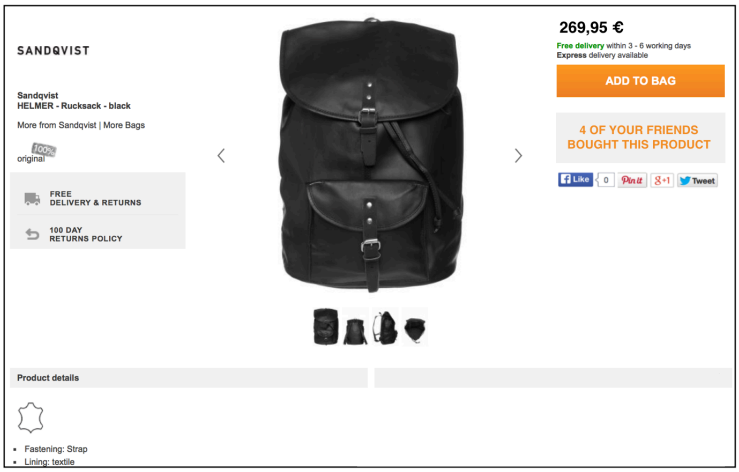
6. *Recommendations* by *customers* on a *high* priced product



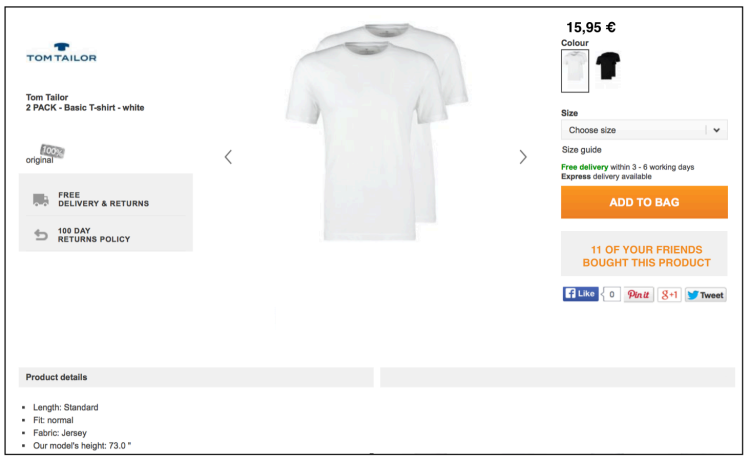
7. *Recommendations* by *friends* on a *low* priced product



8. *Recommendations* by customers on a low priced product



9. Recommendation by *friends* on a *high* priced product



10. *Reviews* by *friends* on a *low priced* product

Appendices 3: Questions for Main Survey Design

Welcome text	Thank you for taking the time to participate in this study. The focus of the study is social influence bias in e-commerce and the estimated timeframe is to 5-7 minutes, but please take your time to answer all the questions as thorough as you can.
1. Do you shop online? (if yes the respondent is directed to Q2, if no/I don't know the respondent is directed to Q6**)**	• Yes • No • I don't know
2. How often do you shop online?	• More than 7 times a month • 4 - 6 times a month • 1 - 3 times a month • Less than once a month • Less than once a year • I don't know
3. Why do you shop online?	• Convenience • Better selection • Lower prices • Better return policies • Other (Please specify)
4. How much do you shop for (on average**)**	• More than 5000 DKK or 650 € • 1000 - 5000 DKK or 160 - 650 € • 500 - 1000 DKK or 60 - 160 € • 200 - 500 DKK or 25 - 60 € • 50 - 200 DKK or 5 - 25 € • Other (Please specify)
5. What do you shop for online?	• Clothing, shoes or accessories • Film, music, books or games • Travel or experiences (eg. concerts) • IT (phone, computer or photo) • Cosmetics, medicine or supplements • Housing or gardening articles • Electronics or appliances • Groceries, food or drinks • Sports equipment • Hobby- or office supplies • Other (Please specify)
6. - 10. On the following pages you will be presented with a series of different images from two major e-Commerce companies. We ask you to examine all the information (price, product, ratings, recommendations etc.) shown on each image before answering the subsequent	Five experimental conditions shown in each survey • How much do you like this product? • How likely are you to add this product to your wish list? • How likely are you to purchase this product?

question.	
11. Which information on the image above influenced you to give the answers to the preceding question?	• The product • The brand • The price or value • The return policy • The speed or price of delivery • The opinion information (reviews, ratings or recommendations) • Other (Please specify)
12. What is your gender?	• Woman • Man • Other
13. What is your occupation?	• Employed for wages • Self-employed • Out of work and looking for work • Out of work, but not currently looking for work • A homemaker • A student • Retired • Unable to work • Other (Please specify)
14. What is your age?	• Younger than 18 years old • 18-24 years old • 25-34 years old • 35-44 years old • 45-54 years old • 55-64 years old • 65-74 years old • 75 years or older
15. What is your nationality?	(Please specify)
Finished → Submit	We appreciate that you took the time to complete this study. Thank you once again!

Appendix 4: Comments from Pilot Test Respondents

Respondent 1:

- It's a very abstract question you're asking all the time with the images, would you buy/basket/….
- If there's no need for something like this, the urge for necessity is also very hard to just asses and reply if you would buy. None of these articles are really on my 'to buy' list, so assessing the urge to click "YES BUY" is quite distorted.
- If you want to check which UI is most appealing, which layout is more attractive, maybe it's easier to do a comparison from the same articles, where you would be more likely to buy it from? Like the same shirt on a different online shop or so? Or create a fictional shop where you place elements yourself just to see which one could be the element to trigger the purchase?

Respondent 2:
Q1: Rephrase as «Have you ever shopped online»
"Do you" implies a timeframe, which is not specified until later questions.
Q6: Image/text alignment is off. See attached image
Instructions preceding
Q6: "On the following pages you will be presented with a series of different images from two major e-commerce companies. We ask you to examine all the information provided on each image before answering the preceeding question." "Preceeding" means "coming before", better to use "following" or similar.
Q7: Text content should be below image, similar problem to Q6. See attached image
Q8: Same issue as Q6
Q10: Would leave "other" as just a choice, without a text box where you need to elaborate further. If you identify as a different gender than purely male/femal eyou might be sensitive to this.
Q11: "Occupation" refers to the specific type of work done. I don't have a better suggestion, unfortunately

Going through the questions it is also a bit unclear on what criteria the user is supposed to rate the pages.

Respondent 3:
Den eneste kommentar jeg umiddelbart har er, at det er svært at finde ud af, om man skal forholde sig til det konkrete eksempler eller bare den måde det er præsenteret på. Og så er det måske bedre på at skrive på dansk ift. formuleringer osv. Medmindre du skal have fat i mange udlændinge eller den skal være engelsk af anden grund.

Eventuelt kan du smide et kommentarlink ind til sidst i besvarelsen - så man ikke skal ind på mail. Nogle af kategorierne under beskæftigelse er lidt svære at tolke (de engelske betegnelser for hjemmegående, lønmodtager)

Respondent 4:
It does not work on iPad

Respondent 5:

- In Abhängigkeit von dem statistische Verfahren, dass sie anwenden will, sollte sie nochmal genau auf die Clusterung der Antworten achten.
- (image) Hier könnte man noch erwähnen, dass multi-selection möglich ist. (genau wie bei einigen der Fragen danach)
- (image) Frage 4 und 5 kommen ja direkt nacheinander – in Abhängigkeit von den Warengruppe, die ich kaufe variiert mein Shopping Wert aber extrem. Z.B. kaufe ich auf Amazon oft Bücher, die 10 Euro kosten je Transaktion und auf Zalando Kleidung mit einem durchschnittlichen Warenwert von 120 Euro.
- Eventuell sollte man das Ganze als Matrix anlegen (wenn es nicht zu schwer für die Analyse ist) ansonsten wird es für den Befragten unheimlich schwer die Zahl einigermaßen richtig abzuschätzen. Mit einer Matrix gibt man den Befragten Halt (Achtung, diese darf nicht zu groß sein, dass gilt als inconvenient für Befragte und wird von Wissenschaftlern als nicht allzu gutes Design bewertet bzw. es besteht die Gefahr, dass die befragten die Umfrage dann abbrechen.)
- (image) Hier passt was nicht mit der Überblende – rechts kann man noch das „Examining" lesen – gilt auch für die anderen Bilder.
- Auf der Startseite könnte sie noch ein bild von sich einfügen auf der sie bereits ihre Mailadresse und das Unilogo einbindet – das wirkt meist noch seriöser und weckt Vertrauen and der Umfrage teilzunehmen.
- Eventuell könnte man auch die Fragen so anlegen, dass die Teilnehmer nicht so stark nach unten scrollen müssen – muss man mal testen.

Appendix 5: Comments from Main Survey

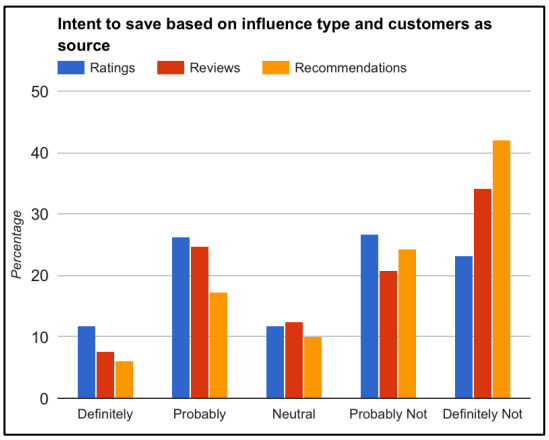
Friends (all)	• Not enough info on product • Product and price • I already own this product • The abundance of information; i was overwhelmed • Nothing • I almost never buy clothes online, and I almost never wear white t-shirts. • Friends buying a piece of clothing would not influence me; I would rather trust a greater group of anonymous strangers. My friends are mostly idiots. Haha :D	• Its not ecofriendly • Description more precise • Not enough information to know if I would like it - why did you hide the reviews? • 11 of your friends bought this product • Boring website • I am a woman I don't wear men's clothes! • Great layout. Easy access/overview of information • No reviews, would not buy • Like the product, hate the price.
Reviews (all)	• I have a vague interest in something called "Interstellar" but I see no information on what it's actually about • I don't need this product. • I would not buy film in dvd but as digital download. • I don't buy dvd's anymore • I'd just bootleg it • I've read reviews of that movie. They certainly don't match the customer reviews and are more impartial than those given by people who have bought the movie. • The abundance of information; i was overwhelmed	• Who buys movies anymore • Not enough info • All of the above • The specs of the TV. 120hz is insufficient. • The courier who delivers to the area likes to throw boxes. I wouldn't buy anything electronic that has to be shipped. • Again, no actual information. What ports does it have? How many consoles can I plug in at once? • Product and price • Not enough info on product • I already own this product
Recommendations (all)	• Wouldn't shop clothes outside my own country - return shipping is too expensive • The product, the brand, I don't wear shorts Same as the shoes. A branded pair of shorts is a waste of money. • preference • Branded products are overpriced for what you get. • they're ugly • The product, the brand, the price, not enough information (materials) • nothing • I almost never buy clothes online, and I almost never wear white t-shirts. • Friends buying a piece of clothing would not influence me, I would rather trust a greater group of anonymous strangers. My friends are mostly idiots. Haha :D	• description more precise • Not enough information to know if I would like it - why did you hide the reviews? • 11 of your friends bought this product • Boring website • I am a woman I don't wear men's clothes! • The product, the brand, the price, not enough information (materials) • The product, the brand, I don't wear shorts • Great layout. Easy access/overview of information • Like the product, hate the price. • No reviews, would not buy • Its not ecofriendly
Ratings	• I only buy out of print or otherwise hard to get books online. • Description • I already have this book • I already own and have read that book • Messy layout • Nothing • not enough info on product	• Product and price • I already own several Canon lenses, so buying a Nikon body wouldn't be smart. • i have a better one • Electronics and the local delivery service do not mix well. That would be delivered by being thrown at the stairs.

Ratings (all)	• I only buy out of print or otherwise hard to get books online. • description • I already have this book • I already own and have read that book • Messy layout • nothing	• not enough info on product • Product and price • I already own several Canon lenses, so buying a Nikon body wouldn't be smart. • i have a better one • Electronics and the local delivery service do not mix well. That would be delivered by being thrown at the stairs.
Customer + Reviews	• I don't need this product. • I would not buy film in dvd but as digital download. • I don't buy dvd's anymore • I'd just bootleg it • I've read reviews of that movie. They certainly don't match the customer reviews and are more impartial than those given by people who have bought the movie. • Who buys movies anymore	• not enough info • All of the above • The specs of the TV. 120hz is insufficient. • The courier who delivers to the area likes to throw boxes. I wouldn't buy anything electronic that has to be shipped. • Again, no actual information. What ports does it have? How many consoles can I plug in at once?
Customer + Recommendations	• preference • Branded products are overpriced for what you get. • they're ugly • The product, the brand, the price, not enough information (materials)	• Wouldn't shop clothes outside my own country - return shipping is too expensive • Same as the shoes. A branded pair of shorts is a waste of money. • The product, the brand, I don't wear shorts
Customer + Ratings	• Messy layout • nothing • not enough info on product • Product and price • I already own several Canon lenses, so buying a Nikon body wouldn't be smart. • i have a better one	• Electronics and the local delivery service do not mix well. That would be delivered by being thrown at the stairs. • I only buy out of print or otherwise hard to get books online. • description • I already have this book • I already own and have read that book
Friends + Reviews	• not enough info on product • Product and price • I already own this product • the abundance of information; i was overwhelmed • Im white. and have style	• I would never wear them, despite price/brand/opinion • Branded product = overpriced again. • They're ugly • Everything
Friends + Recommendations	• I almost never buy clothes online, and I almost never wear white t-shirts. • Friends buying a piece of clothing would not influence me, I would rather trust a greater group of anonymous strangers. My friends are mostly idiots. Haha :D • Its not ecofriendly • description more precise • No reviews, would not buy	• Not enough information to know if I would like it - why did you hide the reviews? • 11 of your friends bought this product • Boring website • I am a woman I don't wear men's clothes! • Great layout. Easy access/overview of information • Like the product, hate the price.

Appendix 6: Chi-Square tests and Bar Chart Visualizations

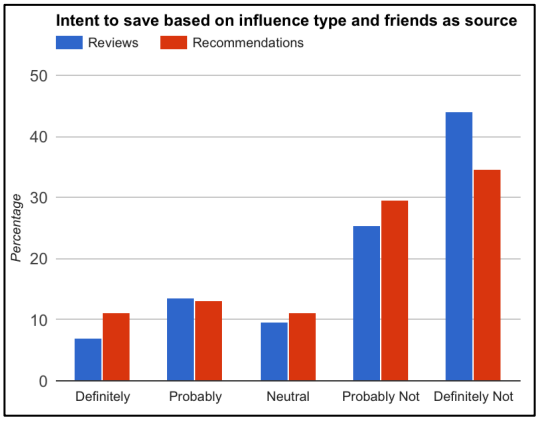
	Intent to save based on influence type and customers as source					
	1	**2**	**3**	**4**	**5**	Subtotals
Ratings	23 *16.61* (2.46)	51 *44.18* (1.05)	23 *22.26* (0.02)	52 *46.51* (0.65)	45 *64.45* (5.87)	**194**
Reviews	15 *16.52* (0.14)	48 *43.95* (0.37)	24 *22.14* (0.16)	40 *46.27* (0.85)	66 *64.11* (0.06)	**193**
Recommen-dations	12 *16.87* (1.40)	34 *44.86* (2.63)	20 *22.60* (0.30)	48 *47.23* (0.01)	83 *65.44* (4.71)	**197**
Subtotals	**50**	**133**	**67**	**140**	**194**	**584**

χ^2 = **20.683**, **df** = **8**, χ^2/**df** = **2.59** , **P(χ^2 > 20.683)** = **0.0080**



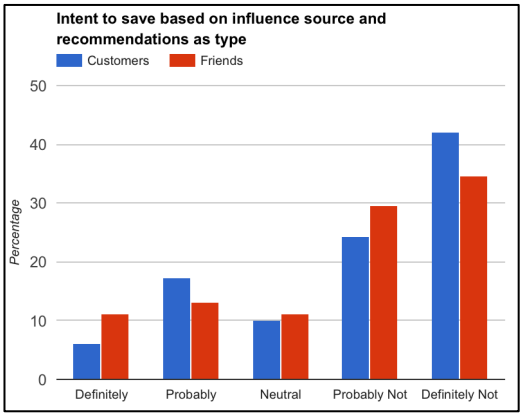
	Intent to save based on influence type and friends as source					
	1	**2**	**3**	**4**	**5**	Subtotals
Review	14 *18.05* (0.91)	27 *26.57* (0.01)	19 *20.55* (0.12)	50 *54.14* (0.32)	87 *77.70* (1.11)	**197**
Recommen-dations	22 *17.95* (0.91)	26 *26.43* (0.01)	22 *20.45* (0.12)	58 *53.86* (0.32)	68 *77.30* (1.12)	**196**
Subtotals	**36**	**53**	**41**	**108**	**155**	**393**

χ^2 = **4.935**, **df** = **4**, χ^2/**df** = **1.23** , **P(χ^2 > 4.935)** = **0.2940**



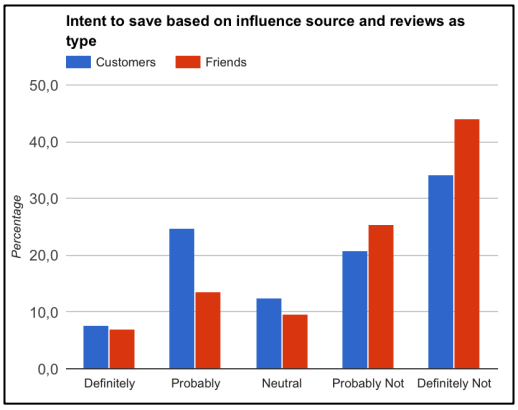
	Intent to save based on influence source and recommendations as type					
	1	**2**	**3**	**4**	**5**	Subtotals
Customers	12 *17.04* (1.49)	34 *30.08* (0.51)	20 *21.05* (0.05)	48 *53.13* (0.50)	83 *75.69* (0.71)	**197**
Friends	22 *16.96* (1.50)	26 *29.92* (0.51)	22 *20.95* (0.05)	58 *52.87* (0.50)	68 *75.31* (0.71)	**196**
Subtotals	**34**	**60**	**42**	**106**	**151**	**393**

χ^2 = **6.534**, **df** = **4**, χ^2/**df** = **1.63** , **P(χ^2 > 6.534)** = **0.1627**



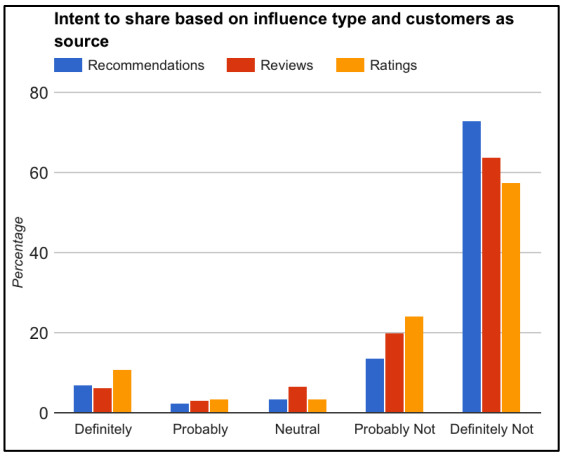
	Intent to save based on influence source and reviews as type					Subtotals
	1	**2**	**3**	**4**	**5**	
Customers	15 *14.35* (0.03)	48 *37.12* (3.19)	24 *21.28* (0.35)	40 *44.54* (0.46)	66 *75.72* (1.25)	**193**
Friends	14 *14.65* (0.03)	27 *37.88* (3.13)	19 *21.72* (0.34)	50 *45.46* (0.45)	87 *77.28* (1.22)	**197**
Subtotals	**29**	**75**	**43**	**90**	**153**	**390**

χ^2 = 10.449, df = 4, χ^2/df = 2.61, P(χ^2 > 10.449) = 0.0335



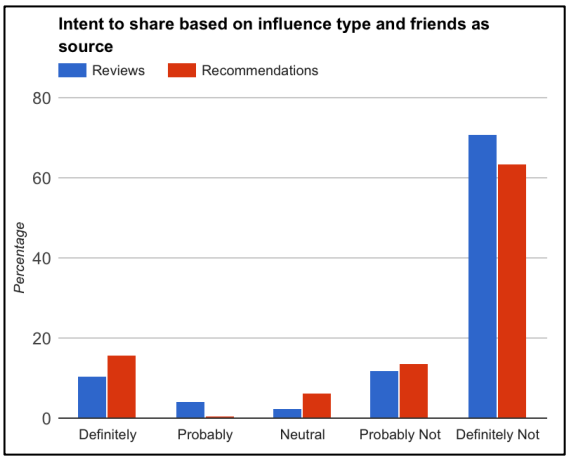
	Intent to share based on influence type and customers as source					Subtotals
	1	**2**	**3**	**4**	**5**	
Ratings	21 *15.42* (2.02)	7 *5.91* (0.20)	7 *8.86* (0.39)	46 *36.75* (2.33)	109 *123.06* (1.61)	**190**
Reviews	12 *15.75* (0.89)	6 *6.03* (0.00)	13 *9.05* (1.73)	39 *37.53* (0.06)	124 *125.65* (0.02)	**194**
Recommen-dations	14 *15.83* (0.21)	5 *6.06* (0.19)	7 *9.09* (0.48)	27 *37.72* (3.05)	142 *126.30* (1.95)	**195**
Subtotals	**47**	**18**	**27**	**112**	**375**	**579**

χ^2 = 15.120, df = 8, χ^2/df = 1.89, P(χ^2 > 15.120) = 0.0569



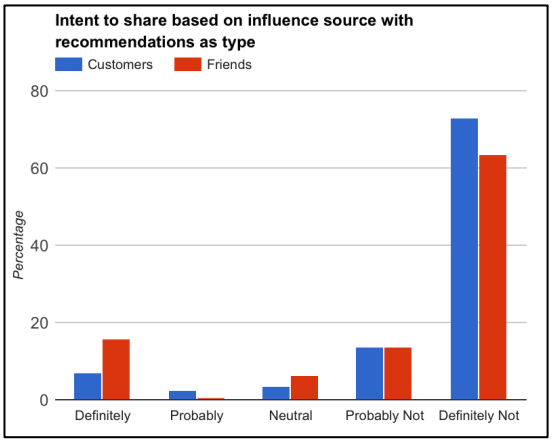
	Intent to share based on influence type and friends as source					Subtotals
	1	**2**	**3**	**4**	**5**	
Ratings	20 *25.13* (1.05)	8 *4.52* (2.67)	5 *8.54* (1.47)	23 *24.63* (0.11)	136 *129.17* (0.36)	**192**
Reviews	30 *24.87* (1.06)	1 *4.48* (2.70)	12 *8.46* (1.49)	26 *24.37* (0.11)	121 *127.83* (0.36)	**190**
Subtotals	**50**	**9**	**17**	**49**	**257**	**382**

χ^2 = 11.376, df = 4, χ^2/df = 2.84, P(χ^2 > 11.376) = 0.0226



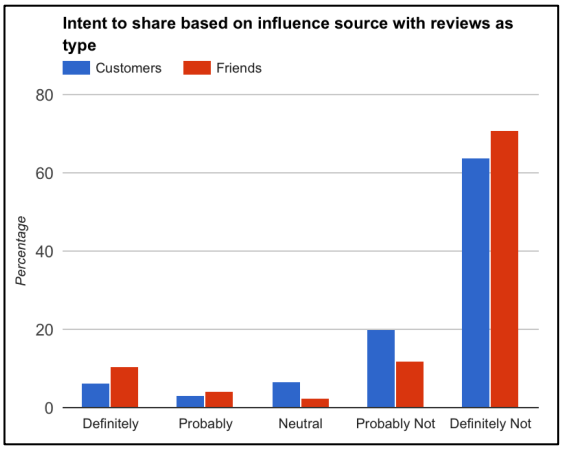
	Intent to share based on influence source with recommendation as type					
	1	**2**	**3**	**4**	**5**	Subtotals
Customers	14 *22.29* (3.08)	5 *3.04* (1.27)	7 *9.62* (0.72)	27 *26.84* (0.00)	142 *133.21* (0.58)	**195**
Friends	30 *21.71* (3.16)	1 *2.96* (1.30)	12 *9.38* (0.73)	26 *26.16* (0.00)	121 *129.79* (0.60)	**190**
Subtotals	**44**	**6**	**19**	**53**	**263**	**385**

χ^2 = 11.433, df = 4, χ^2/df = 2.86, P(χ^2 > 11.433) = 0.0221



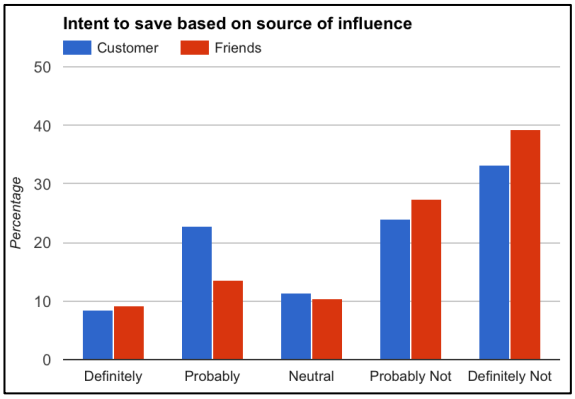
	Intent to share based on influence source with reviews as type					
	1	**2**	**3**	**4**	**5**	Subtotals
Customers	12 *16.08* (1.04)	6 *7.04* (0.15)	13 *9.05* (1.73)	39 *31.16* (1.97)	124 *130.67* (0.34)	**194**
Friends	20 *15.92* (1.05)	8 *6.96* (0.15)	5 *8.95* (1.75)	23 *30.84* (1.99)	136 *129.33* (0.34)	**192**
Subtotals	**32**	**14**	**18**	**62**	**260**	**386**

χ^2 = 10.514, df = 4, χ^2/df = 2.63, P(χ^2 > 10.514) = 0.0326



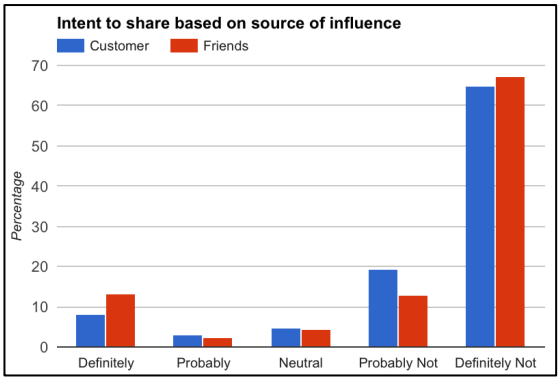
	Intent to save based on source of influence					
	1	**2**	**3**	**4**	**5**	Subtotals
Customers	50 *51.41* (0.04)	133 *111.18* (4.28)	67 *64.56* (0.09)	140 *148.24* (0.46)	194 *208.61* (1.02)	**584**
Friends	36 *34.59* (0.06)	53 *74.82* (6.36)	41 *43.44* (0.14)	108 *99.76* (0.68)	155 *140.39* (1.52)	**393**
Subtotals	**86**	**186**	**108**	**248**	**349**	**977**

χ^2 = 14.654, df = 4, χ^2/df = 3.66, P(χ^2 > 14.654) = 0.005



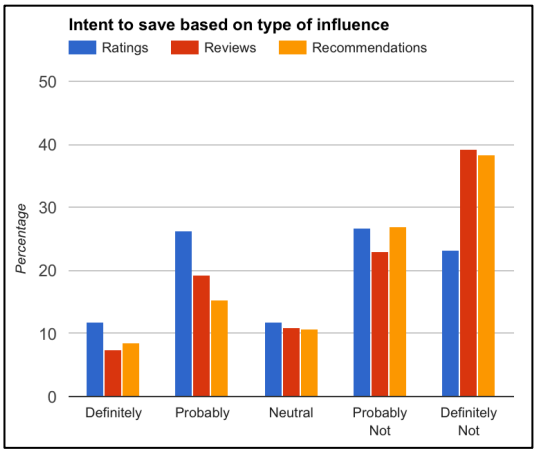
	Intent to share based on source of influence					
	1	**2**	**3**	**4**	**5**	Subtotals
Customers	47 *58.44* (2.24)	18 *16.27* (0.18)	27 *26.51* (0.01)	112 *97.00* (2.32)	375 *380.78* (0.09)	**579**
Friends	50 *38.56* (3.40)	9 *10.73* (0.28)	17 *17.49* (0.01)	49 *64.00* (3.51)	257 *251.22* (0.13)	**382**
Subtotals	**97**	**27**	**44**	**161**	**632**	**961**

χ^2 = 12.177, df = 4, χ^2/df = 3.04 , P(χ^2 > 12.177) = 0.0161



	Intent to save based on type of influence					
	1	**2**	**3**	**4**	**5**	Subtotals
Ratings	23 *17.81* (1.52)	51 *38.51* (4.05)	23 *14.08* (5.65)	52 *51.35* (0.01)	45 *72.26* (10.28)	**194**
Reviews	29 *32.12* (0.30)	75 *69.48* (0.44)	3 *25.40* (19.75)	90 *92.64* (0.08)	153 *130.36* (3.93)	**350**
Recommen-dations	34 *36.07* (0.12)	60 *78.01* (4.16)	42 *28.52* (6.37)	106 *104.02* (0.04)	151 *146.38* (0.15)	**393**
Subtotals	**86**	**186**	**68**	**248**	**349**	**937**

χ^2 = 56.845, df = 8, χ^2/df = 7.11 , P(χ^2 > 56.845) = 0.0000



	Intent to share based on type of influence					
	1	**2**	**3**	**4**	**5**	Subtotals
Ratings	21 *19.18* (0.17)	7 *5.34* (0.52)	7 *8.70* (0.33)	46 *31.83* (6.31)	109 *124.95* (2.04)	**190**
Reviews	32 *38.96* (1.24)	14 *10.84* (0.92)	18 *17.67* (0.01)	62 *64.67* (0.11)	260 *253.85* (0.15)	**386**
Recommen-dations	44 *38.86* (0.68)	6 *10.82* (2.14)	19 *17.63* (0.11)	53 *64.50* (2.05)	263 *253.19* (0.38)	**385**
	97	**27**	**44**	**161**	**632**	**961**

χ^2 = 17.154, df = 8, χ^2/df = 2.14 , P(χ^2 > 17.154) = 0.0285