THE ROLE OF ELECTRONIC WORD-OF-MOUTH IN CONSUMERS’ ONLINE PURCHASE DECISION MAKING: AN EYE-TRACKING STUDY

Master’s Thesis, Marketing

Author: Joel Jokinen

27.12.2016

Supervisors: Matti Leppäniemi
Jarkko Hautala
ABSTRACT

Author
Joel Jokinen

Title
The Role of Electronic Word-of-Mouth in Consumers’ Online Purchase Decision Making: An Eye-Tracking Study

Subject
Marketing

Type of degree
Master’s Thesis

Time of publication
2016

Number of pages
67 + appendices

Abstract
The aim of this study was to shed light on the consumers’ decision making processes in an online environment. The vast amount of information found online and the presence of online peer recommendations has shaped the purchase decision making environment – making it more simple in some situations, more complex in others. This study answers to the need for more research on consumers’ cognitive processes when making purchase decisions, the influence of website design factors towards consumer decision making as well as the social presence of others in online environments. Previously little research has been done on the effects of product ratings toward consumer attention through eye-tracking methodology. Eye-tracking methodology was chosen to overcome the limitations created by using solely self-report methods and projective techniques, such as surveys and interviews, in order to better understand the mental constructs and the behavior of a consumer. A 2 (decision complexity) X 2 (quality of product rating) between-subjects experiment design was employed for this study to assess whether consumers would try to ease cognitively demanding purchase decision making tasks through the use of social heuristics. The data (N=25) was collected through assessing the eye movements of multiple subjects. From the data eye-tracking parameters such as fixation duration, dwell time and the time to first fixation were analyzed through statistical tests. Supporting data was collected through asking the subjects for a brief verbalization of their thought process during the experiment. The results show a significant combined effect of task complexity and product ratings towards the decision making time. No significant combined effect of task complexity and product ratings was found for fixation duration, dwell time and the time to first fixation on the area of interest. A significant main effect was discovered between task complexity and dwell time percentage. Good product ratings were perceived faster than bad product ratings, which as a finding is in line with earlier research. Consumers also seem to be prone to using social heuristics, such as peer-made product ratings, to conform with others during the purchase decision making process, even if the purchase decision is seemingly simple.

Keywords
Consumer behavior, e-commerce, decision making, social comparison, heuristics, eye-tracking, attention

Storage
Jyväskylä School of Business and Economics
FIGURES

FIGURE 1: The Structure of the Study ............................................................. 12
FIGURE 2: The Online Purchase Decision Making Framework ..................... 25
FIGURE 3 The Cyclical Process of Visual Attention ......................................... 35
FIGURE 4: Profile Plot, Time to First Click ...................................................... 50

TABLES

TABLE 1: Key supporting literature for hypotheses ...................................... 31
TABLE 2: The 2x2 Experimental Design ......................................................... 39
TABLE 3: Smartphone Attributes in the Experiment ................................... 42
TABLE 4: Decision Times (Means and Standard Deviations) ......................... 50
TABLE 5: The Main and Combined Effects of Task Complexity and Rating Quality on Eye-Tracking Parameters ......................................................... 52
## CONTENTS

**ABSTRACT**

**FIGURES AND TABLES**

**CONTENT**

1  **INTRODUCTION** ................................................................. 8  
   1.1  Context of the Study ....................................................... 8  
   1.2  Research Problem and Research Questions ............................. 9  
   1.3  Structure of the Study .................................................... 11  

2  **CONSUMER PURCHASE DECISION MAKING PROCESS** ............ 13  
   2.1  Decision Making and Product Choice ................................. 13  
   2.2  Understanding the Complexity of Decisions .......................... 16  
      2.2.1  The Number of Alternatives and the Limitations of Memory  . 16  
      2.2.2  Cognitive Load ...................................................... 17  
   2.3  The Mental Shortcuts of the Mind – Heuristics .................... 18  
   2.4  Attention ................................................................. 20  
   2.5  Consumer Decision Making in the Online Shopping Environment  . 23  

3  **SOCIAL COMPARISON AND ELECTRONIC WORD-OF-MOUTH** ...... 26  
   3.1  The Interest in Social Comparison and Affiliation .................. 26  
   3.2  The Role of Electronic Word-of-Mouth in Consumers’ Purchase  
        Behavior ................................................................. 28  

4  **METHODOLOGY** .................................................................. 32  
   4.1  Eye-tracking Research ..................................................... 32  
   4.2  Eye-tracking Parameters ................................................... 34  
      4.2.1  Attention Revisited .................................................. 34  
      4.2.2  Duration of Fixations ................................................ 36  
      4.2.3  Fixation Density ....................................................... 37  
      4.2.4  Dwell Time ............................................................. 37  
      4.2.5  Time to First Fixation and Path Dependence .................. 38  
   4.3  Experimental Design and Model ......................................... 38  
   4.4  Experimental Procedure ................................................... 40  
   4.5  Validity and Reliability of the Research ............................... 43  
   4.6  Data Analysis ............................................................... 44  

5  **RESULTS** .......................................................................... 45  
   5.1  Demographic and Background Factors ................................. 45  
   5.2  Eye-tracking Results ....................................................... 45  
      5.2.1  Duration of Fixations ................................................ 46  
      5.2.2  Fixation Density ....................................................... 47  
      5.2.3  Dwell Time ............................................................. 47  
      5.2.4  Time to First Click and Decision Time ........................... 49
5.2.5 Time to First Fixation and Path Dependence ..............................51

6 DISCUSSION..............................................................................................53
  6.1 Theoretical Contributions ..................................................................53
  6.2 Managerial Implications ....................................................................55
  6.3 Evaluation of the Research ..................................................................56
  6.4 Limitations of the Research ...............................................................57
  6.5 Directions for Future Research ..........................................................58

REFERENCES..................................................................................................59
Pluribus intentus, minor est ad singula sensus.
1 INTRODUCTION

1.1 Context of the Study

Global B2C e-commerce sales hit a total of over $1.5 trillion in 2014 (eMarketer 2014). The surge of new e-commerce stores and companies has been swift, and the flexibility, convenience and customization e-commerce provides has fundamentally shaped the way people and companies do business (Tezza, Bornia, and Andrade 2011; Luo, Hsieh, and Chiu 2012; Bilgihan and Bujisic 2015). The fierce competition in the e-commerce industry has created a need for companies to develop web sites and online stores that both drive sales through repeated purchases and improve customer loyalty (Srinivasan, Anderson, and Ponnavolu 2002; Chiu, Wang, Fang, and Huang 2014). Hernández, Jiménez, and Martín (2008) as well as Close and Kular-Kinney (2010) state that analyzing consumer behavior in the field of e-commerce is paramount.

Both, academia and companies, agree on the fact that the proliferation of information and product choices available on the internet has drastically changed the consumers’ purchase decision making process (McKinsey & Company 2009; Wu, Shen, and Chang 2014; Zhang, Zhao, Cheung, and Lee 2014). And even more so, academia and companies also agree that encouraging users to generate ratings and reviews online is crucial in the product choice and evaluation processes of the consumer (Microsoft 2013; Flanagin, Metzger, Pure, Markov, and Hartsell 2014). Kim and Srivastava (2007) argue that the incorporation of social influence in the field of e-commerce is becoming more and more important as consumers need the opinions of others to reduce the risk of purchasing a product online.

The consumer decision making process is a thoroughly studied field but the advent of the Internet and the sudden rise of e-commerce has brought new elements to the research. Many of the rules that apply to traditional brick and mortar shopping still apply to online shopping but new areas of interest have risen. As consumers can only process a limited amount of information simultaneously (Miller 1956), the decisions made online have become increasingly complex for consumers. The overwhelming amount of information creates challenges in terms of the cognitive load induced by the display (Sweller 1988) and has given a boost in interest towards the use of heuristics (e.g. Zhang et al. 2014). The active reduction of cognitive efforts plays a major role in the consumer’s rational decision making process (Salant 2011). The exploration of the consumers’ online purchase decision making process is, indeed, gaining more and more academic interest (e.g. De Vries and Pruyn 2007; Tan, Yi, and Chan 2008; Fang 2012; Gao, Zhang, Wang, and Ba 2012; Belanche,
At the same time, humans remain as the social animals we are. Consumers still look up to others to mimic their decisions. (Solomon 2015.) Recent research on social comparison and affiliation among consumers has been done in the areas of, for example, the effect of online recommendations on shopping complexity (e.g. Senecal, Kalczynski, and Nantel 2005), social commerce (e.g. Chen and Shen 2015), and the role social presence in creating customer loyalty (e.g. Cyr, Hassanein, Head, and Ivanov 2007).

Technology has advanced with research methods as well. The consumers’ paths to purchase can be recorded for example through a clickstream analysis or by following the consumers’ eye movements as they follow through with their purchase decision. Recent research on e-commerce that has implemented eye-tracking technology has been done on, for example, information acquisition related to decision making (Shi et al. 2013; Benn et al. 2015), the effect of human brands on consumer decision quality (Chae and Lee 2013), consumers’ decision deliberateness (Huang and Kuo 2011), the effect of product listing pages on consumers’ cognitive load (Schmutz, Roth, Seckler, and Opwis 2010), and consumers’ cognitive processes during online elaboration (Yang 2015).

One could even boldly state that we are living in an era where the consumer decision making process is transforming into something new. The opportunities that lie within e-commerce applications should be explored and academic interest given to research questions related to the transformation of consumer decision making.

1.2 Research Problem and Research Questions

The purpose of this study is to assess the consumers’ path to purchase in an online environment through the lens of cognitive attention towards user-generated product ratings. Often consumers face decision making problems when shopping online and these problems can end up being either complex or simple, in relation to the amount of information and aid given to the consumer. This study will examine whether decision making complexity will influence the need for social comparison and affiliation among consumers in an online environment and furthermore affect their cognitive attention.

This examination is in line with the dire need for more research on the consumer’s cognitive processes when making decisions (Chae and Lee 2013), the influence of website design factors and objects towards consumer decision making (Shi, Wedel, and Pieters 2013; Roth, Tuch, Mekler, Bargas-Avilan, and Opwis 2013), providing consumers with helpful information online (Benn,
Webb, Chang, and Reidy 2015), as well as social presence in online environments (De Vries and Pruyn 2007). Previously little research has been done on the effects of product ratings toward consumer attention through eye-tracking methodology. In addition to providing insights in this area, this study also examines the rather unexplored field of consumers’ attention and cognitive processes during online shopping, in relation to decision making complexity.

Every two years the Marketing Science Institute (MSI) lists top research priorities for marketing given to them by their member companies. These priorities are the ones that member companies consider to drive research initiatives and keep their activities going forward. Naturally, they serve a sign for academia to align their research with the priorities of business done in member companies. For the years 2014-2016 MSI lists “understanding customers and the customer experience” as their top (Tier 1) priority. This essentially includes the question of how technology has shaped consumer behavior. One mentioned area of interest in this top priority is: “How do social media and digital technology change customer experiences and the consumer path to purchase? What are the best ways to model the consumer decision journey? Are other models more appropriate than the decision funnel?”. (MSI Research Priorities 2014-2016.) From this listing it can be concluded that addressing consumer decision making in digital technology related research is paramount.

This study answers to the call for more research on the consumers’ path to purchase as well as research on the cognitive influence and relevancy of different website design factors on consumer decision making. The research questions are as follows:

1) Do consumers rely cognitively on product ratings when making a purchase decision online?
   a. Will purchase decision making complexity influence the need for affiliation among consumers when comparing products online?
   b. Will purchase decision making complexity influence consumers’ use of cognitive heuristics when comparing products online?

2) Is the consumer purchase decision process affected after cognitively processing a product rating?

Due to the nature of the research problem an eye-tracking approach was chosen. It has long been recognized within marketing academia that the limitations created by using solely self-report methods and projective techniques, such as surveys and interviews, need to be overcome when trying to understand the mental constructs and the behavior of a consumer (Haire 1950; Wang and Minor 2008; Chen, Nelson, and Hsu 2015). The eye-tracking methodology that was chosen for this study does just that.
1.3 Structure of the Study

This study has been divided into six chapters. Chapter 1 gives a brief introduction to the underlying theory, explains the purpose of this research and states the research questions. In Chapter 2 the theory behind the consumer decision making process in an online environment is discussed and the key terms used in this study are defined. In addition, in Chapter 2 the cognitive aspects of online consumer behavior, such as heuristics and cognitive attention, are explored. Chapter 3 explains the concepts of social comparison and affiliation and how they appear in the modern day discussions about electronic word-of-mouth (e-WOM). Furthermore, user-generated product ratings are examined in this chapter as a part of the concept of electronic word-of-mouth.

In Chapter 4 the methodological choices made for this study will be explained and argued for. Eye-tracking as a research method will be briefly introduced and the measures used in it explained. In the same chapter the experimental design and procedure of the experiment will be covered in detail. Also, the measures for statistical data analysis will be given.

In Chapter 5 the results from the eye-tracking data and statistical analysis will be discussed. In the concluding Chapter 6 theoretical contributions will be discussed, managerial implications given, and the limitations of the research and directions for future research explained. The structure of the study can be seen also in Figure 1.
FIGURE 1: The Structure of the Study

- Introduction
  - Context of the Study
  - Research Problem and Research Questions
  - Structure of the Study

- Consumer Decision Making in an Online Environment

- Social Comparison and Electronic Word-of-Mouth

- Methodology
  - Eye-tracking Research
  - Eye-tracking Parameters
  - Experimental Design and Model
  - Experimental Procedure
  - Data Analysis

- Results

- Discussion
  - Theoretical Contributions
  - Managerial Implications
  - Evaluation of the Research
  - Limitations of the Research
  - Directions for Future Research
In this chapter the underlying theory of consumer decision making and product choice will be explored. Decision making will also be explained through the lens of task complexity and its effect on the cognitive load of the consumer. In addition, it will be investigated how the cognitive shortcuts consumers make, heuristics, affect decision making, and what drives consumers’ attention during a decision making task. This all will be summarized in the end of the chapter with an outlook on how these aspects come together when consumers are making purchase decisions in an online environment. Based on the underlying research of these concepts, hypotheses will also be provided for the purpose of this study in this chapter.

2.1 Decision Making and Product Choice

People make hundreds of decisions every day (Milosavljevic, Koch, and Rangel 2011). Why do we choose one product over another? Why do some people decide to choose something in an instant while others take their time? These kinds of questions have intrigued researchers of various fields for centuries.

Previous research has two key approaches under which decision making as a human phenomenon can be studied: normative and descriptive. The normative approach investigates the rational and logical nature of decision making, whereas the descriptive approach studies the preferential and belief-based aspects of decision making. (Kahneman and Tversky 1984.) In this study the preferential aspects of decision making are more prominent than the rational aspects. Preferential decision problems usually involve three components: (1) the available alternatives for the consumer, (2) the events or contingencies (and their probabilities) on which the relationship between actions and their outcomes is based on, and (3) the values the consumer associates with the outcome. In an experimental setting these components and the goal statement (e.g. choose the smartphone you most prefer) form the task environment. Naturally this differs from an actual purchase decision setting where the goal statement is not necessarily presented and sometimes the consumer even has to come up with the alternatives themselves. However, it can be said that these components still constitute the basic form of a preferential decision making problem. (Gettys, Pliske, Manning, and Casey 1987; Keller and
Decision making can be seen as a process consisting of three essential subprocesses: information acquisition, information evaluation, and the expression of a decision. To go through this process humans have developed strategies for making decisions and for making choices. Even though there are various different strategies, most of them have some things in common. Decision making strategies often have to resolve value-related conflicts, they may be used separately or combined together, and they can be planned beforehand or built right at the moment when the decision has to be made (e.g. the use of heuristics). Yet, all strategies are different in terms of how much effort the consumer has to put in to use the given strategy and how accurately the outcome of the strategy can be predicted. (Payne et al. 1993.)

Another division for categorizing decisions can be made between risky and riskless decisions. Risky decisions often involve a gamble and certain odds. Riskless decisions in turn are usually transactions where a product or a service is exchanged for, for example, money. (Kahneman and Tversky 1984.) However, the topic of (perceived) risk frequents the recent e-commerce-related discourse (e.g. Kim, Ferrin, and Rao 2008; Belanche et al. 2012; Chiu et al. 2014; Martin et al. 2015), even though Kahneman and Tversky (1984) identify commerce transactions as riskless decisions. Risk in the context of (e-)commerce can be seen as crossing the threshold of trust where the consumer is willing to take a risk in believing that the vendor will act up to expectations (Mayer, Davis, and Schoorman 1995). More so, consumers feel that making purchase decisions online is riskier than making purchase decisions in a store as online they are both spatially and temporally separated from the vendor. Access to information about the purchasable products can be seen to reduce the level of perceived risk. (Tan 1999.)

However, access to product information during the purchase decision making process does not necessarily make things easier for the consumer. It should be noted that the acceptability of a (commercial) transaction for a consumer is often a choice between multi-attribute options. This sometimes creates a value-related conflict, which in turn the consumer tries to solve through the use of heuristics. This means that the consumer needs to set up a mental account to assess the advantages and disadvantages of each option to determine the acceptability of each option. In most cases acceptability is determined by the beneficial relation of advantages and disadvantages. (Kahneman and Tversky 1984.) The chosen strategy for resolving the possible conflict and for making the decision will naturally affect the process as well as the outcome.

In preferential decision problems the formation of preference is of much interest. The consumer can execute the preference formation (PF) through various strategies if necessary. By ‘necessary’ it is meant that the consumer does not always have to form a preference when faced with a choice. They may also have developed a lasting preference for some option, for example in the case of an affect referral. However, if the consumer has a need for preference formation, they have two types of preference formation strategies at their disposal: own-based and other-based preference formation strategies. There also exist hybrids
of these two types of strategies. (Olshavsky 1985.)

Own-based PF strategies include mainly the consumer’s cognitive processes. Of these strategies the most commonly used is decision making. Usually this is done through various decision making strategies and rules as was mentioned earlier. These rules include for example the lexicographic, conjunctive, and expectancy-value based rules. Other own-based PF strategies include the use of cues, judgment, concept identification, learning and reasoning. (Olshavsky 1985; Payne et al. 1993.)

Other-based PF strategies are choice behaviors in which the consumer uses another individual or organization as a surrogate decision maker. Consumers tend to turn to other-based PF strategies when they don’t have a preferred option based on earlier experience or the capacity or willingness to process decision making related information. The most prominent example of an other-based PF strategy is consulting and following a recommendation. It is possible for consumers to use other-based PF strategies for the whole range of the decision making process, including the search for information, the evaluation of options and even carrying out a transaction. (Formisano, Olshavsky, and Tapp 1982; Olshavsky 1985.)

A third division of decision making of interest to this particular study is the division of decision making based on dominated or non-dominated attributes of a product. A dominated alternative is one which is inferior to some other alternative in terms of at least one attribute. On the other hand, a non-dominated alternative is superior to other alternatives on an attribute without being inferior to other alternatives on other attributes simultaneously. Therefore, it can be said that choosing a product objectively would require choosing a combination of both, dominated and non-dominated, attributes in a product. (Häubl and Trifts 2000; Tan et al. 2008.) However, Payne et al. (1993, 88) state that coherent decision making in terms of product choice specifically means not selecting dominated alternatives. This is naturally true in the sense that in a situation when there are, for example, only single-attribute products to choose from, coherent product choices would be directed towards the non-dominated products. In real life this is not the case very often. Objective decisions tend to be about balancing the equation between superior and inferior product attributes.

On the basis of earlier studies, it can be stated that the process nature of decision making involves various different strategies to resolve possible value-related conflicts and to overcome complex decision making tasks. Consumers make decisions and choices based on their earlier experience or the assumed expertise of others. While consumers’ product choice may seem random and emotional, many still undergo a rational process of evaluating the alternatives by their attributes. In order to create simplicity in the consumer purchase decision making process, it is key to understand the meaning of complexity.
2.2 Understanding the Complexity of Decisions

Out of the hundreds of decisions we make every day some are more complex than others. Some may even seem simple to us. It is important to understand what actually constitutes the complexity of a decision making situation.

Formisano et al. (1982, 475) define a difficult task environment as a one with “a large number of alternatives, information on a large number of attributes about each alternative presented in a format that does not lend itself to easy use, and a product or service that is inherently complex”. According to Bennet and Bennet (2004, 290) complexity, in turn, is “the condition of a system, situation, or organization that is integrated with some degree of order, but has too many elements and relationships to understand in simple analytic or logical ways”. Furthermore, the qualities related to a complex decision making situation also include, for example, the diversity of connections, entanglement of patterns, nonlinearity, feedback loops, surprises, uniqueness, and no clear set of alternatives. (Burstein and Holsapple 2008, 5.) Payne et al. (1993, 37–40) also add a temporal dimension to task complexity.

Naturally even the most complex purchase decision situation may not always involve all these qualities mentioned in these definitions but yet they are quite accurate when describing the purchase decisions consumers make online. For the purpose of this study and e-commerce related discourse in general it is worthwhile looking at how the number of alternatives available to the consumer and the related limitations of memory as well as cognitive efforts are involved in the complexity of the decision making process.

2.2.1 The Number of Alternatives and the Limitations of Memory

In previous research related to decision making complexity and choice strategies both Payne (1976) and Olshavsky (1979) have had similar findings. Payne (1976) discovered through information monitoring and protocol analysis techniques that the amount of alternatives determines the choice strategy used by the consumer. Olshavsky in turn (1979) found out that as consumers are presented with more alternatives (i.e. when the decision making situation is made more complex), they switch their choice making strategy from a one-stage, compensatory strategy to a multi-stage strategy. Consumers also tried to simplify the choice making process by assessing and weighing the available information when presented with more alternatives.

Furthermore, Payne (1976) discovered that if the amount of alternatives was increased to 6 and to 12, a two-stage choice strategy was adopted. Here the consumers first screened the alternatives by using a non-compensatory strategy and then used compensatory strategy to evaluate the
rest of the alternatives. Summarized, as the number of the alternatives grows the longer and the more complex the decision making process becomes.

Now, these findings are very much in line with Miller’s (1956) “magical number seven”, also known as Miller’s Law. According to Miller (1956) the human span of immediate memory and absolute judgment limit the ability to receive, process, and memorize information. There actually exists a definite (numeral) limit on a human’s ability to absolutely identify a one-dimensional stimulus variable’s magnitude. The span of our immediate memory and absolute judgment lies approximately around the number seven for one-dimensional judgments. The span of immediate memory is limited by the number of items and the span of absolute judgment by the amount of information. In layman’s terms this means that people can only store approximately seven items in their immediate memory at the same time and process around seven bits of information simultaneously. Naturally people can break longer chains of information into smaller chunks, and in this way overcome the limit of seven. However, this already requires more complex thinking and as Olshavsky (1979) and Payne (1976) found out the strategies used for overcoming situations with more than seven items are far more complex than those with clearly less than seven.

For the purpose of this study these findings and the threshold of the number seven will be used. They will determine what constitutes complex and simple decision making situations in the actual experiment. However, to properly define what constitutes a complex decision making situation, cognitive processing also needs to be taken into account.

2.2.2 Cognitive Load

The consumer’s capacity of cognitive processing is something that should be taken into account when designing functional web stores and creating a path to purchase. Cognitive processing in general consists of two types of activity, information acquisition and internal computation. No matter what kind of a strategy the consumer chooses to use in the purchase decision making process they will always show a pattern of these two. (Russo 1978.) However, the cognitive load, which burdens the consumer in different tasks no matter the strategy, varies according to multiple factors.

While Miller’s (1956) findings about the span of a consumer’s immediate memory and absolute judgment give a clear picture of what constitutes the limitations of a human solving a complex decision making task, they only account for the short-term (immediate) memory. Sweller’s (1988) cognitive load theory (CLT) is concerned with the limitations created by the working memory. The CLT postulates that the cognitive abilities of a human are limited in the sense that they can simultaneously process only a limited amount of entities of information.
According to the CLT there are three types of cognitive load: intrinsic, extraneous, and germane. The intrinsic load means the cognitive load created by the content of the material that is being processed and and the extraneous load is determined by how the material is presented. If the material is hard to process or encode the extraneous cognitive load is larger. Germane load is accumulated through the consolidation of information. In the case of this study, the task complexity creates the intrinsic load and the web store design used in the experiment induces the extraneous load. (Sweller 1988.) According to Wang, Yang, Manlu, Cao, and Ma (2014) the extraneous load can be reduced through clear visual presentation of the material and proper design. It has also been argued by, for example, Payne (1982) and Salant (2011) that consumers themselves also try to reduce their cognitive efforts when solving a problem or making a decision.

Based on earlier research, it is possible to state that decision making complexity can have an influence on consumers’ cognitive processing. It can also be said that consumers try to reduce the amount of cognitive processing through various choice and decision making strategies, trying to find the shortcuts for an easier decision.

### 2.3 The Mental Shortcuts of the Mind – Heuristics

The structure of the consumer purchase decision making process has been studied quite extensively. But what are heuristics and how do they fit into this process?

The rather classical cognitive model of the consumer purchase decision making process includes five stages:

1) Problem recognition
2) Search for information
3) Evaluation of alternatives
4) Choice of product/service
5) Post-choice evaluation of the outcome. (Solomon 2015, 69–80)

In this cognitive decision making model there exists an assumption that the process is linear and sequential and that consumers process information deliberately. (Solomon 2015, 69.) However, there exists previous research, which questions the logicality and rationality of the consumer purchase decision making process on the basis that consumers do not necessarily go through all stages of the process and sometimes make decisions in an instant (Papamichail and Robertson 2008; Karimi, Papamichail, and Holland 2015), in which case the aforementioned process would be quite impossible to go
through wholly. There exists a school of thought that consumers actually often abandon rationality when making purchase decisions and instead opt to take the easiest route that leads to a satisfying decision. These routes are called heuristics. (Solomon 2015, 80–84.)

Heuristics can be defined as “strategies that ignore part of the information, with the goal of making decisions more quickly, frugally, and/or accurately than more complex methods” (Gigerenzer and Gaissmaier, 2011). As consumers try to reduce their cognitive efforts while making decisions (Payne 1982; Salant 2011) heuristics play an important role in this process. Essentially this means that the consumer purchase decision making process can be viewed as two-sided. On the one hand consumers, when thinking about the social world, spend much time and effort in building a decision, while on the other hand they have the possibility of reducing the amount of effort and rely on heuristics. (Moskowitz et al. 1999, 13; Zhang et al. 2014).

Todorov, Chaiken, and Henderson (2002) state that in heuristic information processing “people consider a few informational cues – or even a single informational cue – and form a judgment based on these cues”. This viewpoint differs from Gigerenzer’s and Gaissmaier’s (2011) view in the sense that rather than talking about ignoring information, it focuses on considering certain information, and ultimately making the decision based on that bit of information. However, both definitions agree on the matter that due to heuristics consumers sometimes make decisions based on limited knowledge and do not necessarily take into account the bigger picture.

So why do consumers use heuristics instead of complex, thorough decision making strategies? Would it not be logical that they were to rationalize purchase decisions to make sure that their purchases were advantageous or the best possible? If knowledge is power, why make decisions based on limited knowledge? To gain insight on this dilemma it is worthy to explore the concept of rational decision making briefly.

According to Salant (2011) consumers process information in rational choice tasks based on the identity of the best alternative considered up to that moment. This means that the complexity of the rational choice at hand is almost equal to the amount of viable alternatives. Naturally this in turn means that if the amount of alternatives is great, the choice task becomes cognitively demanding. Consumers may, in this case, resort to a simpler method of resolving the choice problem to reduce or optimize the cognitive costs they have to pay to resolve the problem.

This is also the problem consumers face when shopping, offline or online. Due to the large amount of alternatives, consumers are not able to evaluate all options and their dominant or non-dominant attributes and the shopping task becomes cognitively demanding. Rational choice may not necessarily be an option anymore. Through the use of heuristics consumers are able to direct their attention more swiftly to only a smaller sample of alternatives, and therefore make the purchase decision in an easier fashion. (Wästlund, Otterbring, Gustafsson, and Shams 2015).

Even though heuristics often contain both social and nonsocial information, fully social heuristics can still be identified, for example the
imitation heuristic, the social-circle heuristic, the choosing heuristic, the heuristic of averaging the judgments of others, and the inference heuristic of objectivity. Social heuristics tend to be used in situations where the available information is limited to exploit the so-called wisdom of crowds. (Kruglanski and Mayseless 1990; Hertwig and Herzog 2009; Gigerenzer and Gaissmaier, 2011.)

The imitate-the-majority heuristic means that people tend to observe how others in their reference group behave, and imitate this behavior. When using the social-circle heuristic people search through their own social circles, starting from the closest one to themselves, to determine which alternative to choose. When one alternative arises within a social circle for more times than another, it is chosen. With the choosing heuristic people study quantitative predictions from several advisors using cues for expertise, and choose among these. The averaging heuristic is similar to the choosing heuristic that in both people choose from average quantitative predictions from a few advisors, and in the case of the averaging heuristic, using equal weights. In the inference heuristic of objectivity people consult others who have not been subjected to a possible bias crucially effecting the decision, which the person themselves considered to have been subjected to. (Kruglanski and Mayseless 1990; Hertwig and Herzog 2009.)

As e-commerce websites are riddled with an overwhelming amount of information (Flanagin et al. 2014), on the basis of earlier studies, it can be stated that consumers may indeed resort to using social heuristics in this environment to reduce the cognitive efforts they have to give. There are multiple ways through which e-commerce vendors can creatively apply these cognitive shortcuts into the consumer path to purchase. Directing consumers’ attention during the decision making process may be the key here.

2.4 Attention

The amount of information that a consumer can find on the Internet is mindboggling and can feel even overwhelming. This information-filled environment can create challenges for the consumer, but for businesses as well. Davenport and Beck (2001, 3) go as far as stating that attention has become “the most valuable business currency” in this information era we live in. Davenport and Beck (2001) are not the first ones to emphasize the importance of attention in the consumer decision making process however. Already during the 1970s, Simon (1971, 40; 1978) stated that “a wealth of information creates a poverty of attention” and that “attention is the scarce resource for decision making”. These were wise words, considering that the Internet had not even been invented at
this time. Attention is a universal topic in decision making with no temporal restrictions.

Summarized from earlier research, attention can be defined as the selectivity of perception. Much of earlier research gives high importance to understanding what constitutes the selective attention of a consumer during the decision making process. (Orquin and Loose 2013). As a baseline, one could say that our eyes reflect our attention, meaning that we are generally paying attention to what we are looking at (e.g. Posner 1980). Attention as a visual process will be covered more in detail in Chapter 4.2.1 Attention Revisited.

However, there is more to attention than meets the eye. Even though there exists a close relationship between eye movements and attention, they are separable (Bashinski and Bacharach 1980; Posner 1980). According to the pioneering research by Yarbus (1967) and later on by Posner (1980), the direction of attention can happen through an endogenous (central) control of attention or it can be drawn by peripheral stimuli through exogenous (reflexive) control of attention. They are also called the goal-driven (top-down) and the stimulus-driven (bottom-up) forms of attention, respectively (Orquin and Loose 2013). Directing attention as well as eye movement through external signals requires that the stimulus is of importance to the person (Posner 1980). Posner, Snyder, Davidson (1980) found out that when the correlation between a stimulus of importance to the subject and the foveal location of the eyes is broken, the touch point to attention disappears. In this experiment the subjects detected a bright spot of light faster if their attention had been directed to this spatial location by a cue.

In turn, through their experiments Bashinski and Bacharach (1980) were able to posit that attention can be moved to a potential source of stimulus before the stimulus has actually happened. This means that people can actually move their attention somewhere without moving their eyes. In any research related with the combination of eye movement and cognitive attention it must be taken into account that attention and the foveal structure of the visual system do not necessarily have a straightforwardly causal relationship.

However, in the same study Bashinski and Bacharach (1980) were also able to prove that if the attention of the subject is allocated to a certain spatial location, their visual sensitivity towards that location increases. This was evident after they had placed a locational cue, which made possible for the subjects to temporally shift their attention to that cue, without moving their eyes. In other words, the placement of a cue attracts attention, even though it may not be seen from eye movements.

According to Smith and Ratcliff (2009) attention is in interaction with variables such as visual masks, external noise in the display, and spatial uncertainty. Their integrated theory of attention and decision making in visual signal detection posits that attention controls how a representation of a stimulus forms in the visual short-term memory. The visual short-term memory works in the way that it encodes the outputs of a stimulus in a durable form and preserves it long enough to make a decision. Attention improves the efficiency of this process. So, ultimately what attention does to decision making is that firstly it limits the decision to the stimulus that the consumer is fixated on, and
secondly increases the influence of the information that the consumer is fixated on. (Orquin and Loose 2013).

Even though the goal-driven control of attention is usually stronger than the stimulus-driven control of attention, there are also several factors that enable salient features to affect attention more than the top-down control. These include semantic or contextual cues, attention-based features, representations of objects, and task performance rewards. (Tatler, Hayhoe, Land, and Ballard 2011.) Therefore, in decision making tasks these kinds of salient features may be fixated on, and furthermore influence the decision making process, regardless of the feature’s importance to the decision at hand (Orquin and Loose 2013).

In addition to the visual field and salient features, task relevance is a major driver of attention (Yarbus 1967). This means that when making a decision a consumer will react more preferentially to a stimulus that has high task relevance, possibly ignoring stimuli with low task relevance. The relevancy of the stimulus is naturally for the consumer to decide but it is possible that consumers learn to categorize stimuli into relevant and irrelevant through practice and experience. (Orquin and Loose 2013). Consumers generally tend to assess online information through the use of heuristics (e.g. following the recommendation of an other) (Flanagin et al. 2014) and consumer attention tends to be directed at task-irrelevant stimuli in simple decision making tasks (Wang et al. 2014). This could in an online shopping context, for example, meaning that more experienced online shoppers would define different salient cues as relevant than inexperienced shoppers.

The relationship between attention and working memory is another area of interest. The eye-mind hypothesis posits that what is being fixated on reflects what is being processed (Just and Carpenter 1976). Increases in the working memory load (i.e. increases in task complexity) linearly increase the number or the duration of fixations (Just and Carpenter 1976). Although there exists critique to the linearity of this relationship, it has been generally validated. Complex, cognitively difficult decisions (e.g. decisions with many attribute relationships and dependencies) also cause more intentional re-fixations to lower the demands created for the working memory. (Orquin and Loose 2013.)

On the basis of the relationship between attention and task complexity, and the notion of salient features affecting the consumer decision making process, the first and the second hypotheses can be introduced:

**H1:** During a complex decision making task the subjects will perceive the product rating faster and cognitively process it more than during a simple decision making task, even when objective, non-social means of evaluation are available.

**H2:** During a simple decision making task the subjects will perceive the product rating slower and cognitively process it less than during a complex decision making task.

These hypotheses will be further elaborated on in Chapter 3 due to their multidimensionality.

Attention is only one of the many variables that affect the consumer decision making process. Even though the relationship between visual field and
attention is close, consumers can shift their attention to stimuli of their interest without moving their eyes. Understanding the process of shifting attention plays an important role in the information-rich environment the Internet offers. Getting the consumers’ attention at the right time, on to the right location, may lead to interesting results that are exhibited in the online shopping behavior of consumers.

2.5 Consumer Decision Making in the Online Shopping Environment

The amount of online transactions has globally increased within the recent years. Even though the field is growing and the business is booming, academia has found that there are also some impediments that influence the consumers’ online purchase decision making process. (Chae and Lee 2013). These include for example the lack of social interaction, the absence of personal consultation (Barlow, Siddiqui, and Mannion 2004) and the lack of trust in products as well as the companies who sell them (Kim et al. 2008). Understanding how to overcome these inhibitors will enable companies to create a simpler path to purchase and an improved online shopping experience.

As has been previously stated in this study, the consumer decision making process is influenced by a plethora of factors. The number of alternatives, the limitations of the immediate as well as the working memory, the cognitive load induced by the situation, the proneness to the use of heuristics, and attention all affect what the purchase decision to be made will be. Do these same rules work when consumers are making decisions in an online shopping environment?

To begin with, as was stated earlier in this study, consumers feel that purchase decision made over e-commerce web sites are risky (e.g. De Vries and Pruyn 2007; Belanche et al. 2012; Chiu et al. 2014; Martin et al. 2015). Due to the spatial and temporal separation from the vendor, consumers feel that making purchase decisions online is riskier than making purchase decisions in a brick and mortar store (Tan 1999.) The feeling of risk can be reduced during the purchase decision making process for example through accessibility, visibility, and ease of use (Martin et al. 2015). In addition, the incorporation of social influence in web stores is paramount as consumers crave for the opinions of others to reduce the risk of purchasing a product online (Kim and Srivastava 2007).

In addition to risk assessment, consumers also need to face the dilemma of making good, satisfying decisions. Sometimes if consumers do not have previous experience about the products they are planning on purchasing, they turn to other-based decision making strategies. As was stated earlier in this
study, the most prominent other-based decision making strategy for consumers is following a recommendation. (Olshavsky 1985). By including elements that implement the use of other-based decision making strategies such as social influence, online reviews (Cyr et al. 2007), or online interactivity (Fang 2012), consumers are able to find extra cues to help them in the purchase decision making process. In this way e-commerce vendors can positively enhance the online shopping experience, and furthermore affect the consumers’ decision quality. (Fang 2012).

Another important aspect to take into account in the consumer purchase decision making process is the number of alternatives and the limitations of human memory. As it is now known consumers can only process approximately seven chunks of information simultaneously (Miller 1956). While online consumers are exposed to rather complex shopping tasks in terms of the available information and this may make processing and responding to this information more difficult. This in turn influences the path to purchase so that consumers may only consider a limited amount of alternatives or they may choose to ignore vital information. Due to this the ultimate purchase decision may not end up being the most optimal one. (Tan et al. 2008; Gao, Zhang, Wang, and Ba 2012).

When online consumers are faced with intrinsic, extraneous, and germane cognitive load (Sweller 1988; Wang et al. 2014). Even though the amount of information induces more cognitive load, the most informative websites are of the ones that capture the attention of the consumers. Website complexity affects consumer decision making, depending on task complexity. In simple decision making tasks attention tends to be focused on the task and it does not necessarily spill to irrelevant elements on the web site. (Wang et al. 2014.)

Earlier in this study it was explained how consumers employ heuristic decision making strategies when evaluating purchase alternatives. With the internet being such an information-rich environment consumers have to come up with ways to cope with the sometimes excruciating amount of product evaluation-related informational cues in web stores. Due to this, consumers tend to often opt for the use of cognitive heuristics to evaluate the credibility of online information and to make decisions based on that (Wolf and Muhanna 2011; Flanagin et al. 2014.)

In today’s consumer psychology, understanding the final purchase decision is not enough. The whole process with its perceptual and cognitive aspects needs to be understood. One way to achieve this goal is to analyze the eye movements of consumers while they go shopping online – how they behave and what do they attend to. (Chae and Lee 2013).

Summarized, the consumer purchase decision making process is dynamic and flexible and consumers adapt their ways of reacting to different decisions tasks on the basis of multiple factors (Karimi et al. 2015). The framework suggested by Karimi et al. (2015) (Figure 2) depicts how the consumer decision making process has turned into something that is not linear nor sequential, but rather a process with loops and constant re-evaluation.

In the next chapter the second perspective of this study will be
explored. The concepts of social comparison, affiliation and electronic word-of-mouth will be explored. As was already seen in this chapter, social elements truly play a part in the consumer decision making process and the online path to purchase.

FIGURE 2: The Online Purchase Decision Making Framework (Karimi, Papamichail, and Holland 2015)
3 SOCIAL COMPARISON AND ELECTRONIC WORD-OF-MOUTH

In this chapter the phenomena of social comparison and affiliation will be explored. These concepts will be introduced through the scope of consumer marketing as well as the psychology of heuristics. The modern day forms of social comparison, particularly electronic word-of-mouth (e-WOM) and user-generated product ratings, will also be explained. Based on the underlying research of these concepts, hypotheses will also be provided for the purpose of this study in this chapter.

3.1 The Interest in Social Comparison and Affiliation

It is not often that a consumer ends up making a purchase decision solely by themselves. The behavior of others determines one’s behavior. The effect of interpersonal influence in consumer decision making can be seen in effect for example when advertisements depict products being used in social situations or by famous people. Consumers in general are susceptible to interpersonal influence. (Bearden, Netemeyer, and Teel 1989.) The theories of social comparison and affiliation explain the consumers’ urge to validate themselves socially and fall under the influence of the opinions of others.

The concept of social comparison is based on Festinger’s (1954) influential conceptual framework, the theory of social comparison processes. The theory defines a person’s need for social comparison as “a drive to evaluate his opinions and his abilities” (Festinger 1954, 118). The theory posits that a person’s (i.e. a consumer’s) behavior is affected by the person’s cognition (opinions and beliefs) of the situation they are in and what they are capable of doing in that given situation. After all, if the person has made an incorrect assessment of the situation, it may lead to disastrous results. Therefore, the drive to evaluate one’s opinions and capabilities arises.

The influence of others reaches even deeper into the decision making of consumers than what Festinger (1954) originally suggested. His work has been further elaborated on by for example Schachter (1959) and Goethals and Darley (1977). Whereas Festinger (1954) was concerned with how people compare their opinions and abilities, Schachter (1959) was concerned with emotional comparisons. In turn, Goethals and Nelson (1973) studied the comparison of values and beliefs. These studies and their theoretical
contribution – even though they are quite old – are still widely used in modern day marketing studies, related to for example electronic word-of-mouth (e.g. Brown, Broderick, and Lee 2007) and consumers’ online choice making (e.g. Zhu and Huberman 2014).

According to Schachter (1959, 5) consumers do not just have a drive for evaluating their opinions and abilities but rather a “general drive for cognitive clarity”. He posits that a consumer’s cognitive needs are what constitutes their affiliative needs. The link between a consumer’s cognitive processing and the need for affiliation is prominent.

We need to ask what triggers this internal drive in people to compare themselves to others? Many previous studies have confirmed that social comparison is often associated with situations of uncertainty, stress, novelty, and change (e.g. Festinger 1954; Taylor, Buunk, and Aspinwall 1990; Wills and Suls 1991; Gibbons and Buunk 1999). These kinds of situations will momentarily increase the amount of social comparison behaviors. Schachter (1959) also posits that emotional distress will increase the need for affiliation.

In addition to situational factors there are also naturally intrinsic factors that relate to a person’s individual attributes. These include for example the fear of invalidity (Kruglanski and Mayseless 1987), the need for confirmation or cognitive structure, the tendency to search for similarly minded others (Kruglanski and Mayseless 1987), and the need for cognitive closure (Kruglanski, Webster, and Klem 1993). For example, the study by Kruglanski and Mayseless (1987) found out that in a situation where a subject feels high fear of invalidity, they compare themselves more with disagreeing others. In turn, if the subject is categorized to have a high need for self-confirmation or cognitive structure, they compare themselves more with agreeing others.

However, there are also limitations on how people compare themselves socially. If the person that is being compared to is too different from the person that is making the comparison, no comparison will happen in terms of assessing one’s opinions and abilities. Therefore, the closer one feels to another person, the more likely they are to make the comparison to their opinion. (Festinger 1954.) Now, what is also interesting is that Festinger’s (1954) theory also posits that people engage in social comparison behaviors only when objective, non-social means of evaluation are not present. Yet, for example Olshavsky and Granbois (1979, 98) argue that many product choices are solely based on non-decision making rules (i.e. not objective nor social), such as the heuristics of “conformity to group norms, imitation of others” and following “recommendations from personal or non-personal sources”. Therefore, it is hypothesized in this study that the use of social heuristics and the need for social comparison will override the objective, non-social decision making means of the consumer (see Hypothesis 1, page 22). This also provides ground for Hypothesis 3. Based on the fact that consumers tend to use social heuristics when they have little or no previous knowledge about the situation (Gigerenzer and Gaissmaier 2011), and on the fact that consumers tend to compare their opinions socially when faced with uncertainty (e.g. Festinger 1954; Taylor et al. 1990; Suls and Wills 1991; Gibbons and Buunk 1999), and that consumers generally try to minimize cognitive efforts when making decisions (Salant 2011;
Wang et al. 2014), it can be hypothesized that:

**H3: During a complex decision making task the subjects will perceive the good product rating faster and cognitively process it more than a bad product rating.**

In the consumer context it is primary to discuss the concepts of social comparison and affiliation as the consumer susceptibility to interpersonal influence (McGuire 1968; Bearden et al. 1989). The concept is defined as “the need to identify with or enhance one's image in the opinion of significant others through the acquisition and use of products and brands, the willingness to conform to the expectations of others regarding purchase decisions, and/or the tendency to learn about products and services by observing others or seeking information from others.” (Bearden et al. 1989, 473).

Based on earlier research, Bearden and Rose (1990) state that there are four sources of social comparison information in the consumer context: (1) behavioral cues, (2) explicit announcements by members of important reference groups of the consumption of a product, (3) the social rewards and punishments within the important reference groups, and (4) the possible reactions of the group to the consumer’s purchase behavior.

Even though the underlying theories may seem like remnants from ancient times to some, the concepts of social comparison, affiliation, and interpersonal influence are still relevant in modern marketing and human-computer interaction studies. Consumers’ susceptibility to interpersonal influence has been studied for example from the viewpoints of offline as well as online purchase intentions (Shukla 2011; Chen, Teng, Yu, and Yu 2016), online customer loyalty (Cyr et al. 2007), consumer habits and purchase-related behavior (Lee 2016; Koller, Floh, Zauner, and Rusch 2013), the salesman’s influence (Sun, Tai, and Tsai 2009), and even online game choices (Lee 2015). It can be stated that grounding an e-commerce-related study on the basis of the theories of the human drive to evaluate one’s opinions and abilities is more than worthwhile. The modern day studies of the social influence of others in purchase decision making relate more closely, however, to the concept of electronic word-of-mouth and the immense use of user-generated product reviews and ratings.

### 3.2 The Role of Electronic Word-of-Mouth in Consumers’ Purchase Behavior

Emphasizing the importance of word-of-mouth as a marketing method is more than justified. Previous studies have found out that consumers perceive word-of-mouth information as more trustworthy than that of traditional media and advertising, and that word-of-mouth affects consumers’ purchase decisions (e.g.
With the advent and rise of the Internet word-of-mouth has been taken to a whole another level as consumers are now able to post, search for, and share word-of-mouth information online in a fast and convenient fashion. (Cheung and Thadani 2012.)

The concept of word-of-mouth (WOM) has been traditionally defined as “a process of personal influence, in which communications between a communicator and a receiver influence consumer purchase decision” (Cheung and Thadani 2012). Mazzarol, Sweeney, and Soutar (2007) state that the persuasiveness of WOM information derives from the lack of biased selling intentions. Consumers consider WOM information to be more credible and trustworthy than traditional marketing information, and therefore pay more attention to it (Brown et al. 2007). On the basis of these earlier studies, Hennig-Thurau, Gwinner, Walsh, and Gremler (2004) define electronic word-of-mouth (e-WOM) as “any positive or negative statement made by potential, actual, or former customers about a product or company, which is made available to a multitude of people and institutions via the Internet.”

Chatterjee (2001) states that the dynamics of traditional WOM are applicable to e-WOM but that they are also different in the sense of the modes of communications, the volume of information and their commercial focus. In turn, Cheung and Lee (2012) identify four points of difference, which distinguish e-WOM from traditional WOM: (1) the speed of diffusion and an exponential potential for scalability (see also De Valck, Van Bruggen, and Wierenga 2009), (2) the persistence and accessibility of communications, (3) the measurability of communications, and (4) the receiver has to judge the credibility of the information based on various cues, such as online ratings. In addition, Cheung and Thadani (2012) state that the presentation format of e-WOM has made it easier to observe than traditional WOM.

Even though most e-WOM is in a text-based format (Cheung and Lee 2012), user-generated product reviews and ratings are one major form in which e-WOM appears on the Internet (Chatterjee 2001). The interest towards their use in commercial applications has grown significantly as of late. Consumers use these ratings to assess the credibility of commercial product information and to reduce the risk of purchasing a product. It has been found out that the higher the average of the product ratings is, the perceived product quality as well as the consumer’s purchase intention increases. (Flanagin et al. 2014.) Gupta and Harris (2010) have also found out that e-WOM increases the time a consumer considers the recommended product. Evaluation of products online is generally affected a lot by e-WOM (Doh and Hwang 2009).

Some online vendors encourage their customers to give reviews of their products as studies have found out that consumers would rather buy products that have been reviewed by (trusted) peers than those which have not. Many vendors do not, however, use product ratings as effectively as they could. (Kim and Srivastava 2007; Flanagin et al. 2014.)

Wolf and Muhanna (2011) found out from their study of seller’s product ratings that consumers associated strong ratings with a higher level of trust. Flanagin et al. (2014) had similar findings in their study of user-generated ratings and their relationship with perceived product quality. In addition, Park
et al. (2007) found out that the quality of an online rating influences consumers’ purchase intention. In addition, Gupta’s and Harris’ (2010) findings support the idea that a recommended product (i.e. a one with good ratings) would get more attention than one that is not recommended (i.e. a one with bad ratings or no ratings at all). Simplified, this means that the higher the product rating is, the more consumers trust it and there more they are likely to shift their attention towards it. From this it is possible to postulate Hypothesis 4:

H4: The presentation of a good product rating influences the subject’s cognitive and decision making process more than the presentation of a bad product rating, regardless of task complexity.

As Flanagan et al. (2014) stated there are still many e-commerce vendors who do not use user-generated product ratings as effectively as they could. As was earlier mentioned, according to Chae and Lee (2013), one way understand the consumer online store experience is to analyze the eye movements of consumers while they go shopping online. This study aims at combining these two needs for research, derived from the fields of business as well as academia. In the next chapter the methodology of conducting this study will be presented to shed light on these issues.
### TABLE 1: Key supporting literature for hypotheses

<table>
<thead>
<tr>
<th>Hypotheses</th>
<th>Key supporting literature</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>H1:</strong> During a complex decision making task the subjects will perceive</td>
<td>Festinger 1954; Schachter 1959; Just and Carpenter 1976; Taylor, Buunk, and Aspinwall 1990; Wills and Suls 1991; Gibbons &amp; Buunk, 1999; Gigerenzer and Gaissmaier, 2011; Wolf and Muhanna 2011; Flanagin et al. 2014; Wang et al. 2014</td>
</tr>
<tr>
<td>the product rating faster and cognitively process it more than during a</td>
<td></td>
</tr>
<tr>
<td>simple decision making task, even when objective, non-social means of</td>
<td></td>
</tr>
<tr>
<td>evaluation are available.</td>
<td></td>
</tr>
<tr>
<td><strong>H2:</strong> During a simple decision making task the subjects will perceive</td>
<td>Festinger 1954; Schachter 1959; Just and Carpenter 1976; Taylor, Buunk, and Aspinwall 1990; Wills and Suls 1991; Gibbons &amp; Buunk, 1999; Wolf and Muhanna 2011; Flanagin et al. 2014; Wang et al. 2014</td>
</tr>
<tr>
<td>the product rating slower and cognitively process it less than during a</td>
<td></td>
</tr>
<tr>
<td>complex decision making task.</td>
<td></td>
</tr>
<tr>
<td><strong>H3:</strong> During a complex decision making task the subjects will perceive</td>
<td>Festinger 1954; Taylor, Buunk, and Aspinwall 1990; Wills and Suls 1991; Gibbons and Buunk 1999; Gigerenzer and Gaissmaier, 2011; Salant 2011; Wang et al. 2014</td>
</tr>
<tr>
<td>the good product rating faster and cognitively process it more than a</td>
<td></td>
</tr>
<tr>
<td>bad product rating.</td>
<td></td>
</tr>
<tr>
<td><strong>H4:</strong> The presentation of a good product rating influences the subject’s</td>
<td>Park et al. 2007; Gupta and Harris 2010; Wolf and Muhanna 2011; Flanagin et al. 2014</td>
</tr>
<tr>
<td>cognitive and decision making process more than the presentation of a</td>
<td></td>
</tr>
<tr>
<td>bad product rating, regardless of task complexity.</td>
<td></td>
</tr>
</tbody>
</table>
4 METHODOLOGY

In this chapter the methodology used for this study will be reviewed and argued for. A brief history and the applicability of eye-tracking research methods and its respective measures and parameters will be given. In this chapter the appropriateness of the experimental research strategy for the purpose of this study will be argued for, the 2x2 experimental design will be introduced and the experimental procedure will be explained in detail. In addition, the methods for data analysis, statistical as well as eye-tracking based, will be given.

4.1 Eye-tracking Research

As of recently, psychophysiological research techniques have gained ground within marketing research to reveal consumers’ perceptions and emotional reactions related with marketing stimuli (Wang and Minor 2008). These research techniques also make avoiding the problems of self-report methods possible. Marketing research has long been troubled by problems such as consumers not fully articulating their thoughts or emotions or providing answers that they consider socially acceptable. By using psychophysiological research techniques, such as eye-tracking methodology, researchers are able to collect biometric data and measure the actual perception and behavior of consumers. (Haire 1950; Wang and Minor 2008; Chen et al. 2015.) Eye-tracking as a psychophysiological method provides the researcher with an accurate picture of the consumer’s decision processes while maintaining a naturalistic take on data collection. (Huang and Kuo 2011).

According to Duchowski (2003, 3) there are two general motives for studying eye movements. They provide a psychological view on the attentional behavior of humans and on the other hand they give a physiological perspective on what is causing this behavior on a neural level. Why is this interesting for consumer and marketing research? What perspective can we gain from eye movements when exploring the consumer path to purchase or designing the layout of our web stores?

Eye fixations are the most efficient way for us humans in general to acquire information from our surroundings (Russo 1978). Furthermore, in most human-computer interactions the primal touch point between the two parties is naturally the eye (Jacob 1995). As e-commerce applications and interfaces are solely based on human-computer interaction, the same primal touch point
applies. Seeing as how e-commerce web sites are a very visual environment for shopping, it is more than relevant that their quality should be researched with methodology that implements the visual gazing behavior of consumers (Romano, Bergstrom, and Schall 2014).

The roots of eye-tracking as a research method can be traced back all the way down to the late 1800s but even the modern consumer and marketing research of eye movements traces back to the 1970s (Russo 1978; Rayner 1998; Wang and Minor 2008). Eye-tracking is a research method where the eye fixations and movements of the subject are studied. This kind of research allows the researcher to know where, how long and in what sequence were the subject’s eyes fixated on in real-time. What makes this interesting in terms of consumer and marketing research is, however, the underlying eye-mind hypothesis, which posits that eye fixations reflect and reveal the consumer’s cognitive processes and mental operations. The areas where the subject fixates their eyes reflects attention given to that area and in turn the length of that fixation reflects the amount of cognitive processing the consumer has to make for the given area of interest. (Just and Carpenter 1976; Poole and Ball 2005.)

Recent e-commerce-related research by using eye-tracking techniques has been done on, for example, information acquisition related to decision making (Benn et al. 2015; Shi et al. 2013), the effect of human brands (Chae and Lee 2013), recommender interfaces (Castagnos and Pu 2010; Chen and Pu 2010, 2014), controlling visual attention (Ho 2014), consumer impulsivity (Huang and Kuo 2012), consumers’ decision deliberateness (Huang and Kuo 2011), product listing pages (Schmutz et al. 2010), and the application of the Elaboration Likelihood Model (Yang 2015). Eye-tracking research related to internet advertising and the concept of banner blindness has also been increasingly popular in the field of e-commerce and web advertising (for example, Teixeira, Wedel, and Pieters 2012; Porta, Ravarelli, and Spaghetti 2013; Luo et al. 2014; Resnick and Albert 2014; Hernández-Méndez and Muñoz-Leiva 2015).

Shi et al. (2013) emphasize that we can learn much of the consumer’s decision making and product choice processes by analyzing eye movement data. There has also been some previous eye-tracking research on decision-making and product choice due to the benefits of this research method (for example, Glöckner and Herbold 2011; Khushaba, Wise, Kodagoda, Louviere, Kahn, and Townsend 2013; Shi et al. 2013; Venkatraman, Payne, and Huettel 2014; Vu, Tu, and Duerrschmid 2015; Waechter, Sütterlin, and Siegrist 2015). However, there is not exactly an abundance of decision-making related eye-tracking research that has been done for e-commerce purposes, even though the field would much benefit from it.
4.2 Eye-tracking Parameters

There are two main measurements for eye movements used in eye-tracking research: fixations and saccades. Fixations are moments when the eyes are quite stationary and saccades are rapid eye movements that occur between fixations, and are used to move the focus of the eye to a new visual location. Encoding information takes place during a fixation whereas no cognitive encoding occurs during saccades. Therefore, it is not possible to tell much of how a web site object affects a subject through analyzing saccades. Regressive saccades can tell us, however, if the subject had trouble encoding some of the information. (Duchowski 2003, 44, 48; Poole and Ball 2005.) There are also three smaller types of eye movements: nystagmus, drifts, and microsaccades (Rayner 1998). Due to the nature of this experiment, analysis will be conducted through the use of fixation parameters. Saccades as well as the small movement parameters will be excluded.

4.2.1 Attention Revisited

Attention is an integral part of eye-tracking research. Even though attention as a concept was earlier in this study explained through the scope of decision making, it is vital to understand the optical side as well. This way interpreting the data gained from the subject’s fixations and saccades is possible to encode.

When we look straight ahead our visual field consists of three regions: foveal, parafoveal, and peripheral. Only the most central 2 ° of our visual field (the fovea) is very good in terms of acuity. 5 ° on both sides of a fixation are still moderate in terms of acuity (the parafovea), but the rest, our peripheral visual field, is already quite poor. In order to see the object clearly we have to move our eyes, so that the fovea overlaps with the object we want to see. (Rayner 1998.)

It is actually possible to move our attention towards something without moving our eyes to fixate on it (Posner 1980). However, most often the objects that stimulate us are so complex that it is just easier to move our eyes, so we can focus our attention on them (Peiyuan and Kowler 1992). Therefore, in complex information processing tasks the locus of attention and the location of the eyes go hand in hand (Posner 1980; Rayner 1998).

There exist two major viewpoints on how visual attention is constituted. They do not exclude one another and have formed the basis for modern concepts of visual attention. The “what” (James 1981) reflects the foveal part of our visual attention and the “where” (Von Helmholtz 1925) reflects the parafoveal part. (Duchowski 2003, 5.)
According to Von Helmholtz (1925) when we direct our visual attention to objects in the peripheral field, it reflects our willingness to inspect these objects. The spatial location, or in essence the “where” of visual attention, is therefore his primary concern. James (1981) on the other hand considered attention to be a similar construct to imagination, anticipation and thought. The “what” of attention means his association of identity, meaning and expectation with the focus of our attention. (Duchowski 2003, 5.)

On the basis of these two grounding viewpoints a human’s visual attention can be seen as a cyclical process (as depicted in Figure), consisting of the following steps:

1) A stimulus is first seen through peripheral vision. This resorts into other interesting objects appearing in the peripheral field of vision. They direct the vision towards them, to engage the fovea.
2) The initial stimulus loses the foveal location of the eyes. The eyes repositioned to the new object of interest.
3) The eyes fixate on the new object of interest with the foveal location of the eyes being on that object. Attention is directed to perceive the object in high resolution. (Duchowski 2003, 12)

![Figure 3: The Cyclical Process of Visual Attention (Adapted from Duchowski 2003, 12)](image)

Understanding this cyclical process of visual attention enables us to interpret the fixations and saccades a consumer is making when browsing in a web store.
4.2.2 Duration of Fixations

Fixations are states where the retina is moderately stably focused on one area of interest (Duchowski 2003, 48). The duration of fixations reflects the complexity, or the simplicity, of information integration on that given area of interest. Therefore, fixations also reflect the cognitive processes behind directing one’s visual attention towards an object of interest. (Glöckner and Herbold 2011.) The (mean) duration of fixations can be used to measure these cognitive processes and is calculated by dividing the fixation times by the fixation count (Waechter et al. 2015).

In usability and e-commerce related research the duration of fixations is much of interest. Much of the previous research done on the connection of the duration of fixations and subject’s cognitive processing (e.g. Velichkovsky 1999; Velichkovsky, Rothert, Kopf, Donoher, and Joos 2002; Cowen, Ball, and Delin 2002; Horstmann, Ahlgrimm, and Glöckner 2009; Glöckner and Herbold 2011) argues that an increased mean length of a fixation’s duration means an increased level of cognitive processing. Naturally, in turn, if the mean duration of fixations is shorter, this would mean that the information has been easier to process. Poole, Ball, and Phillips (2004) suggest on the basis of this notion that fixating on an object of interest means that the object is more important and interesting to the subject. Duchowski (2003, 169), in turn, states that during previous problem solving related applications of eye-tracking research the more difficult aspects of the problem solving process have generated longer durations of fixations.

There exists, however, an opposing school of thought to this phenomenon as well. For example, Just and Carpenter (1976) have argued that while the duration of fixations may reflect the cognitive processes of the subject, it is not possible to know whether this means that the subject is having difficulties in processing the information or whether the area that the subject is fixated on is more of interest. However, for the purpose of this study, this argument does not provide a challenge as the area of interest is simple to process and its symbolic meaning well known to the subjects.

The mean fixations of durations cannot be classified definitively into categories but the thresholds used depend on the experiment (Holmqvist, Nyström, Andersson, Dewhurst, Jarodzka, and Van De Weijer 2011). Velichkovsky et al. (2002) state that in previous research it has been found out that in a visual search-and-compare task, such as the tasks given to the subjects in this study, there exist two phases of fixation durations. The first phase, approximately from 150 to 250 milliseconds, functions just in terms of spatial density and configuration but in the second phase the fixation duration is increased to 500 milliseconds. At this stage the salient features of the physical area of interest do not control the fixation duration anymore, but instead the complexity of the decision controls it. In this study it is of interest whether the subject focuses more on the product rating during a complex (or a simple) decision making task, so these thresholds can be used as guidelines for analysis.
Overall, for the purpose of this study the mean fixation of durations is a valid parameter for measuring the cognitive processing and information integration undergone by the subject (i.e. Hypotheses 1 and Hypothesis 3).

4.2.3 Fixation Density

Another eye-tracking measure related to fixations that was used in this study is the number of fixations, also known as fixation density. Fixation density means that the subject’s fixations are focused on a small area of interest at the time of the fixation. This indicates that the attention is focused on that area and the search for information is efficient. In turn, if fixations are spread evenly on the display the search for information can be considered to be inefficient. (Cowen et al. 2002; Poole and Ball 2005).

Fixation density is also an indicator of attention. A higher amount of fixations on a target area of interest indicates higher importance of this area to the subject (Rayner 1998; Poole et al. 2004; Waechter et al. 2015). In an encoding task a high fixation density indicates interest towards the object at hand (Jacob and Karn 2003; Chae and Lee 2013).

Naturally the visual features of the display are in a crucial role when directing the subject’s attention to a small area of interest (Duchowski 2003, 154; Waechter et al. 2015). This is why fixation density has been used as a metric in, for example, advertisement-related eye-tracking research to measure which information has subject processed within a single advertisement (Venkatraman et al. 2014; Venkatraman et al. 2015). Thus, this parameter also suits the purpose of this study as the intention is to find out if the subjects focus their attention on the product rating, which is a fairly small area within the web store displayed to the subjects (i.e. Hypotheses 2 and Hypothesis 4).

4.2.4 Dwell Time

Dwell time is another eye-tracking measure that was used for the purpose of this study. Dwell time, or in simpler terms gaze or gaze duration, is the sum of all fixation durations and saccades on the area of interest. As longer dwell times generally mean more cognitive processing, this metric can be used for comparing the attention given to, for example, two separate areas of interest. Dwell time can also be used to measure the subject’s anticipation of something if their gaze dwells on some area before an expected incident occurs. (Henderson & Hollingworth 1999; Hauland 2003; Poole & Ball 2005; Waechter et al. 2015.)

In this study, dwell time was used to measure the cognitive processing done by the subjects with regards to the area of interest. The
combination of the fixation density and dwell time reflect the depth of the information integration process within the area of interest. The longer the dwell time and the fewer the fixations, the more the area of interest is processed. (Horstmann et al. 2009; Venkatraman et al. 2014; Venkatraman, Dimoka, Pavlou, Vo, Hampton, Bollinger, Hershfield, Ishihara, and Winer 2015.)

4.2.5 Time to First Fixation and Path Dependence

Scan path means the sequence of fixations and how they are spatially arranged. Scan path is an indicator of the efficiency of the layout of the elements used in the interface the subject is looking at. (Jacob and Karn 2003.) In this study it was investigated how fast the product ratings will be perceived by the subjects, so taking account the scan path is relevant. A clear, straightforward scan path would indicate a fast observation of the target area of interest, whereas a dispersed scan path would indicate a longer search.

According to Venkatraman et al. (2014) the path dependence of the subject reflects judgment and choice behavior. They posit that the subject’s use of a heuristic choice strategy with the aim of maximizing utility and minimizing the cognitive load will show up as path dependence.

The time used by the subject to first fixate on the target area of interest is also a valid parameter for measuring Hypothesis 1, 2 and 3, seeing as how a specific search target exists research-wise (Jacob and Karn 2003). Naturally the subjects themselves did not necessarily have the target area of interest as their search target but the time to the first fixation on the (target) area of interest is of much interest to the researcher in this study.

4.3 Experimental Design and Model

For this study it is most appropriate to use an experimental research strategy. An experimental research offers the best possible solution for testing the causality between two variables. In addition, there is little or no ambiguity about the direction of the causality, which is a challenge that sometimes occurs when conducting other forms of quantitative research, such as survey-based research and related structural equation modelling. Another reason for the appropriateness of the experimental research strategy is that the study is investigating a micro-level phenomenon. (Bryman and Bell 2007;
Causality can be identified when two phenomena are related, they have temporal precedence, and a third variable, different from the independent variable, is not able to cause the effect. In the case of social sciences, the information gained from experimental research can be considered the most valuable. (Metsämuuronen 2005, 6-7.) This research can be considered as an explanatory study as its aim is to find causal relationships between the independent and dependent variables (Hirsjärvi, Remes, and Sajavaara 2005, 129).

A 2 (decision complexity) X 2 (good vs. bad product rating) between-subjects experiment design was employed for this study. The decision making task complexity (independent variable) was divided into complex and simple tasks and the division was made by using Miller’s (1956) aforementioned findings of a human immediate memory capacity and limitations of absolute judgment. In addition, adding information about the products increases the intrinsic cognitive load induced by the task (Sweller 1988), therefore transforming the tasks with more product information into more complex ones. Leuthold’s, Schmutz’s, Bargas-Avila’s, Tuch’s, and Opwis’ (2011) study was also imitated in the sense that by adding criteria to a task, the task complexity increases. The goodness of the product rating was numerical. A rating had a minimum of 0.5 stars and maximum of 5.0 stars, where the amount of stars represented the goodness of the rating. In between the minimum and the maximum of stars, the amount of stars was changing but only so that the number was either a full number or a half (e.g. 3 stars or 3.5 stars). 3.5 stars was used as the separator between good and bad ratings. This resulted in a total of four experimental conditions:

**TABLE 2: The 2x2 Experimental Design**

| CG = Complex Decision Making Task + Good Rating | SG = Simple Decision Making Task + Good Rating |
| CB = Complex Decision Making Task + Bad Rating | SB = Simple Decision Making Task + Bad Rating |

Other stimuli material presented to the subject, barring the changes in decision making task complexity and the goodness of the product rating, were identical in all four conditions. The order in which the subjects were posed to each condition varied, so that there were no order effects that could have affected the results. There were no changes in the factor levels during or after the experiment.
For the experiment the iView X Hi-Speed eye-tracker (SMI, Germany) was used. Monocular gaze and pupil data was collected from each subject’s right eye movements. The specialized software integrated with the hardware generated x- and y-coordinates to determine the subjects’ gaze points on a monitor screen in front of them. Experiment Centre 2.0 (SMI, Germany) was used to create and run the experiment. The room used for the experiment was free of any extraneous disturbances and the lighting in the room was also set not to disturb the subjects. The researcher remained in the room while the subjects were undergoing the experiment but did not interfere once instructions had been given.

All subjects read and signed a consent form, which asked for demographic information, such as age, sex, optical problems, previous use of e-commerce websites, and previous use of mobile phones, as well as for their permission that their gaze behavior would be recorded and used for the purpose of this study. The subjects also acknowledged that the data would be treated anonymously and if at some stage of the experiment they felt uncomfortable, they could quit and leave the room.

Before beginning each experiment session with each subject the eye-tracker had to be calibrated personally for each of them in order to ensure good data quality. Each subject was seated leaning on the eye-tracker facing the monitor. Their chin rest and their seating position were adjusted accordingly. Each subject was instructed to remain as motionless as possible throughout the experiment. After a suitable position for each subject had been found, the researcher explained the experimental procedure. The subjects were instructed to take their time with the experiment as time was not an issue.

The eye-tracker was calibrated using the eye-tracker’s own calibration software to ensure that each subject was able to view the monitor so that the data would be satisfactory. The calibration was acceptable when the maximum deviation did not exceed $x < 1,0^\circ$ and $y < 1,0^\circ$. The calibration was repeated with each subject until successful, up to five times per subject. The calibration was also done again halfway through the experiment to ensure that data quality remained consistent through the experiment.

Following the calibration, the instructions were explained to each subject on the monitor as follows:

“You are shopping for a smartphone in a web store. Find a smartphone that suits your preferences. After you have decided which smartphone best suits your preferences, click “Purchase” to select and purchase the given smartphone. After you have purchased the smartphone, a new task will appear on the screen. When you have completed all tasks, it will be informed to you on the screen. At that moment you can quit the experiment.”

The test was explained in the native language of each subject (Finnish) to ensure
full understanding. The researcher also incentivized the subjects to behave as realistically as possible during the experiment. The incentive was two movie tickets per subject. Two pre-tasks were given to each subject before the actual tasks to confirm that they had understood their task for the experiment.

Each subject received a total of 36 tasks on a fictional e-commerce web site of which 18 tasks were complex tasks and 18 tasks were simple tasks. Each task was given only once to each subject, so there were no same tasks twice per subject. The shopping tasks were instructed to be executed in the order they appear, so the subjects were not allowed to skip any of the tasks. The 18 complex tasks had 9 bad ratings and 9 good ratings and the 18 simple tasks also had 9 bad ratings and 9 good ratings. Even though the amount of ratings can vary a lot without affecting the outcome (Flanagin et al. 2014), it was decided that this should not alter the decision making process in this experiment, so the amount of given ratings shown to the subject varied only a little. It was also decided that the score of the product ratings could vary a little to create a more realistic feeling. There were no written, verbal or pictorial cues determining which of the tasks were simple or complex. Other elements on the fictional web site were kept at minimum to avoid unnecessary stimuli, yet maintaining the resemblance to an actual web store to create a sense of reality. The web store layout can be found in Appendix 1.

In each task the subjects were presented a choice between 4 smartphones. Smartphones were used because of relevancy to student samples and because currently smartphones are such a broad product category where choices can be complex due to the market’s sheer variety as well as product attributes (Tan et al. 2008).

Each alternative was described with either under 7 or over 7 product attributes, depending on task complexity (Miller 1956). There was, however, a clear distinction between complex and simple tasks in terms of the amount of product attributes (e.g. for simple tasks only 4 product attributes and for complex tasks 9 attributes), so that individual differences would not affect the decision making processes’ complexity. The amount of smartphones shown as choices to each subject was kept at four, so it would be as simple as possible to every subject.

Two constant attributes were given to the smartphones but the rest of the attributes were varied.

---

For example, Gupta and Harris (2010) used laptops due to their relevancy to student samples and because they could be seen as experience products. They do not mention the basis of relevancy but it can be assumed that students are familiar with consumer electronics, such as laptops and smartphones.
TABLE 3: Smartphone Attributes in the Experiment

<table>
<thead>
<tr>
<th>Constant attributes</th>
<th>Varied attributes (number of levels in parentheses)</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Operating System (Android)</td>
<td>• Price (3)</td>
</tr>
<tr>
<td>• Size</td>
<td>• Battery (2)</td>
</tr>
<tr>
<td></td>
<td>• Camera (2)</td>
</tr>
<tr>
<td></td>
<td>• Screen (2)</td>
</tr>
<tr>
<td></td>
<td>• Memory (3)</td>
</tr>
<tr>
<td></td>
<td>• Navigation (2)</td>
</tr>
<tr>
<td></td>
<td>• Network Connection (2)</td>
</tr>
<tr>
<td></td>
<td>• Durability (2)</td>
</tr>
</tbody>
</table>

In order to ensure that the decisions each subject made were objective, barring the influence of the product rating, one out of the four smartphone models in each task was non-dominated. Being non-dominated means that the product was always generally the best alternative given all the product attributes. A non-dominated product could have some weaker attributes than a dominated product but was overall the single best choice for the subject to make. The choice of a non-dominated product indicates objective decision quality. The use of non-dominated (vs dominated) products also rules out the effects of brand preference. Each brand was given one non-dominated product, though they were mutually dominated. (Payne et al. 1993, 88; Häubl and Trifts 2000; Tan et al. 2008.) Objective decision quality is relevant for this study to ensure that decisions are made through thorough evaluation of alternatives, so that every alternative receives the subjects’ attention. The subjects’ biased attention may have lead to unnecessary variables affecting the experimental setting.

The rest of the dominated products were given attributes based on the levels of various attributes. This way no model that would appear in each task would be the same and the decision making process would not be affected by unnecessary factors, and the non-dominated products would stay non-dominant. (Häubl and Trifts 2000.) For the price of the product only 3 price levels were used, so that individual price sensitivity will not affect the results. The price levels were also quite close to one another². The product attributions and descriptions can be found in Appendix 2 (Product Descriptions: Smart Phones).

The layout of the fictional web store was given a realistic design. However, the resolution of the human eye’s retina declines faster vertically than horizontally, resulting in a horizontally wider visual field. This means that humans often have left-to-right eye movements when scanning or searching for information and cues. (Gilchrist and Harvey 2006). Hence, the product order on

---

² Tan et al. (2008) use price as a constant attribute. However, for the purposes of this study, it was decided that it would influence the feeling of reality in the experimental situation too much if price was kept the same. After all, in a real-life shopping situation it is very unlikely that smartphones would cost the same amount of money.
the screen was randomized in the sense that the non-dominated alternative would not always be on the left side of the screen. This way the left-to-right eye movement did not influence decision making and there was no bias in the results.

The web store was not however a fully functioning web store but pictures taken from a fictional web store, created specifically for the purpose of this study. The subjects were instructed to solely click on the blue "Purchase"-buttons to choose a product and move on to the next task. Technically they could have clicked anywhere on the screen, which would have moved them to the next task, endangering data quality. None of the subjects clicked anything other than the "Purchase"-buttons prematurely.

The experimental setting was done by each subject in two parts with the second calibration in between. After the first part was done it was informed to the subject on the screen. This was done also after the second part, after which the subject was asked to briefly describe in writing what kind of a strategy had they had when making purchase decisions. This was done to gain additional cognitive insight from their decision making. After describing their strategy, the subjects were given their movie tickets and escorted out of the experiment room.

In the next chapter it will be explained how the data from the experiment will be analyzed and what methods will be used for the analysis.

4.5 Validity and Reliability of the Research

According to Metsämuuronen (2005, 12–14) there are several threats for validity when conducting experimental research. As the aim of an experimental study is to control the variables, and through this gain validity, the independent variable should be the cause of the effect or result. In the case of this study the independent variable can be seen as a cause of for the desired effect. In the case of this study the historical, maturation and regression effects that threaten the validity of research do not affect the results of the study, even though a control group was not used for this study.

A manipulation check was, however, conducted to ensure the reliability of the actual experiment. Without a manipulation check drawing any ultimate conclusions is difficult (Metsämuuronen 2005, 19). In the manipulation check two participants were instructed in a similar fashion as the subjects of the actual experiment and were given a similar set of tasks to execute in the fictional web store. After the tasks had been executed by the participant, they were asked two questions. Firstly, did the participant feel that decision making was complex during the tasks that were supposed to be complex (decision making complexity), and secondly, whether the participant felt that they would
have needed more assistance during the tasks (social comparison). The participants felt that the complex tasks were, indeed, more complex than the simple tasks and felt that they would have needed more information, i.e. help, to execute the tasks more successfully than with the information at their disposal.

4.6 Data Analysis

In order for valid data to be collected from the experiment, areas of interest (AOIs) had to be set for each task. The AOIs for each task were set using SMI BeGaze 3.5 software (SMI, Germany). AOIs were set for the pictures of the smartphones, the product descriptions, the product ratings and the Purchase - button. The AOIs were set by forming a rectangular area around each object. For the purpose of this study only the data from the product rating AOIs was analyzed.

The eye-tracking data was collected by the data analysis program SMI BeGaze 3.5, which worked in junction with Experiment Centre 2.0 (SMI, Germany). The data was then exported in Microsoft Excel, which was also used for the data analysis to gain averages, variances and standard deviations of various eye-tracking metrics, such as dwell time, decision time and fixations. IBM SPSS Statistics was used to conduct statistical tests to the data for further insights and levels of significance in the effects of variables towards one another. Two-way within-subjects repeated measures ANOVAs were calculated for each parameter.

The results of the data analysis will be covered and explained in the following chapters.
5 RESULTS

In this chapter the results of the study will be presented. Firstly, the demographic and other background factors of the experiment will be covered. Secondly, the statistical data and the results from the eye-tracking data will be investigated.

5.1 Demographic and Background Factors

Data was collected from a total of 25 subjects (N=25). 56% of the subjects were male. The average age of the subjects was 25.25 years. None of the subjects reported to have any optical problems.

When asked about the previous experience of using web stores on a scale from 1 to 6, the average of 2.64 indicated that the subjects used a web store at least on a monthly basis. When asked about the previous experience using smartphones on a scale from 1 to 6, the average of 1.04 indicated that the subjects used a smart phone multiple times in a day. When asked about the knowledge the subjects felt they had about the technical attributes of a smartphone on a scale from 1 to 4, the average of 1.68 indicated that the subjects felt they knew the technical attributes of smartphones well. The scales and the form used to collect this information can be found in Appendix 3.

5.2 Eye-tracking Results

In this chapter the results that deal with the eye-tracking parameters will be looked into in detail. All parameters are given descriptive measures and statistical tests were conducted to further investigate the effects between different variables.
5.2.1 Duration of Fixations

In this study mean fixation duration was used as the metric for the duration of fixations. Mean fixation duration is fixation time divided by fixation count. Mean fixation duration reflects the cognitive processes behind directing one’s visual attention towards an AOI (in the case of this study, the product ratings). (Glöckner and Herbold 2011.)

Even though separate AOI areas were set for each product rating respectively in each task, for the purpose of this study it was not relevant which individual rating the subject was concentrated on. Therefore, for the analysis of mean fixation duration the four separate AOIs in each task were treated as a single AOI.

According to Hypothesis 1, during a complex decision making task the subjects will cognitively process the product ratings more than during a simple decision making task, even when objective, non-social means of evaluation are available. And vice versa, according to Hypothesis 2, during a simple decision making task the subjects will cognitively process the product ratings less than during a complex decision making task. Mean fixation duration was used to measure cognitive processing as longer mean fixation duration reflects more cognitive processing.

For simple decision making tasks the mean fixation duration in milliseconds was [ms]: M [SD] = 595,82 [290,56]. For complex decision making tasks the mean fixation duration in milliseconds was [ms]: M [SD] = 608,07 [308,74]. In order to explore the within-subjects variance a two-way repeated measures ANOVA was conducted with fixation duration as the dependent variable and task complexity and the quality of the product rating as the independent variables. The interaction of these variables did not reach a significant level, F(1,25) = 0,489 (p = 0,491). The combined effect of task complexity and quality of ratings does not have a significant impact on the mean fixation duration on the AOI. When the independent variables were explored individually, the interaction between fixation duration and task complexity did not form a significant main effect, F(1,25) = 0,384 (p = 0,541) nor did the interaction between fixation duration and rating quality, F(1,25) = 0,207 (p = 0,653).

Overall, it can be said that the comparison of means indicates that during complex decision making tasks the subjects cognitively processed the product ratings more than in simple decision making tasks. However, as no statistical significance was discovered, these results cannot be taken as definitive.
5.2.2 Fixation Density

For the purpose of the time to first fixation metric the product ratings were treated as one common AOI rather than having one product rating represent one AOI. As per Hypothesis 1, 2, and 3 it is not of interest which product rating is being cognitively processed but, in general, how much are the product ratings processed.

To measure fixation density, the fixation count metric was used, which shows numerically how many times a subject has fixated on an AOI. For all tasks the mean fixation count was $M \ [SD] = 14,18 \ [11,90]$. For complex decision making tasks the mean fixation count was $M \ [SD] = 13,74 \ [11,59]$. For simple decision making tasks the fixation count was $M \ [SD] = 14,62 \ [12,23]$. For complex decision making tasks with a bad rating the mean fixation count was $M \ [SD] = 13,54 \ [10,97]$. For the complex decision making tasks with a good rating the mean fixation count was $M \ [SD] = 13,94 \ [12,23]$. For the simple decision making tasks with a bad rating the fixation count was $M \ [SD] = 14,58 \ [13,19]$. For the simple decision making tasks with a good rating the fixation count was $M \ [SD] = 14,66 \ [11,27]$. In order to explore the within-subjects variance a two-way repeated measures ANOVA was conducted with fixation density as the dependent variable and task complexity and the quality of the product rating as the independent variables. The interaction of these variables did not reach a significant level, $F(1,25) = 0,018 \ (p = 0,894)$. The combined effect of task complexity and quality of ratings does not have a significant impact on the fixation density on the AOI. When the independent variables were explored individually, the interaction between fixation density and task complexity did not form a significant main effect, $F(1,25) = 0,367 \ (p = 0,550)$ nor did the interaction between fixation density and rating quality, $F(1,25) = 0,018 \ (p = 0,894)$.

The comparison of the means for the fixation density parameter would indicate that the subjects unexpectedly cognitively processed the product ratings more during simple decision making tasks, as opposed to what was suggested by the hypotheses. It needs to be acknowledged, though, that as no statistical significance was found, these results are not definitive. Again, if solely the means were to be compared, it was found that in line with Hypothesis 4, regardless of task complexity, good product ratings were cognitively processed more than bad ratings. However, no statistical significance was found, so these results are not conclusive.

5.2.3 Dwell Time

Dwell time was used in this study to measure the general cognitive processing a subject gave to the AOI. Hypotheses 1 and 2 postulated that during a complex decision making task the subjects would process the product ratings more than
during a simple decision making task. For the purpose of this parameter the product ratings were treated as one common AOI rather than having one product rating represent one AOI. As per Hypothesis 1 and 2 it is not of interest which product rating is being cognitively processed but, in general, how much are the product ratings processed.

The mean net dwell time in milliseconds was [ms]: $M \ [SD] = 3452.97 \ [3053.04]$ whereas the mean dwell time in milliseconds was [ms]: $M \ [SD] = 2995.34 \ [2768.82]$. The mean net dwell time in milliseconds for simple tasks was [ms]: $M \ [SD] = 3453.66 \ [2985.87]$. The mean net dwell time in milliseconds for complex tasks was [ms]: $M \ [SD] = 3452.82 \ [3126.43]$. The mean net dwell time in milliseconds for tasks with good ratings was [ms]: $M \ [SD] = 3399.29 \ [3022.13]$. The mean net dwell time in milliseconds for tasks with bad ratings was [ms]: $M \ [SD] = 3506.64 \ [3090.44]$.

In order to explore the within-subjects variance a two-way repeated measures ANOVA was conducted with the mean net dwell time as the dependent variable and task complexity and the quality of the product rating as the independent variables. The interaction of these variables did not reach a significant level, $F(1,25) = 0.127 \ (p = 0.725)$. The combined effect of task complexity and quality of ratings does not have a significant impact on the dwell time on the AOI. When the independent variables were explored individually, the interaction between the mean net dwell time and task complexity did not form a significant main effect, $F(1,25) = 0.009 \ (p = 0.926)$ nor did the interaction between the mean net dwell time and rating quality, $F(1,25) = 0.026 \ (p = 0.874)$.

Whereas these results are not statistically significant and seem to be inconsistent it is worthwhile taking into account that the product rating AOIs covered a rather small area of the whole picture presented to the subjects. Therefore, studying the dwell time percentages is crucial for acquiring meaningful results.

The mean net dwell time percentage was $M \ [SD] = 17.78 \ [11.85]$ whereas the mean dwell time percentage was $M \ [SD] = 15.39 \ [10.87]$. In order to explore the within-subjects variance a two-way repeated measures ANOVA was conducted with the mean net dwell time percentage as the dependent variable and task complexity and the quality of the product rating as the independent variables. The interaction of these variables did not reach a significant level, $F(1,25) = 0.030 \ (p = 0.864)$. The combined effect of task complexity and quality of ratings does not have a significant impact on the mean net dwell time percentage. When the independent variables were explored individually, the interaction between the mean net dwell time percentage and task complexity formed a significant main effect, $F(1,25) = 58.112 \ (p = 0.000)$. The interaction between the mean net dwell time percentage and rating quality did not form a significant main effect, $F(1,25) = 0.001 \ (p = 0.978)$. 
5.2.4 Time to First Click and Decision Time

The time to the first click (ie. decision time) was used in this study to measure the general cognitive processing a subject gave to the whole task. Hypothesis 4 postulated that the presentation of a good product rating alters the subject’s cognitive and decision making process more than the presentation of a bad product rating, regardless of task complexity. The time each subject spends on each decision would therefore be altered by the presentation of a product rating.

As the subjects in the experiment were under no time pressure or restrictions, the time used to make a decision varied among different tasks as well as different subjects. The decision time in seconds was [s]: M [SD] = 18,28 [13,01]. The decision time for complex decision making tasks is seconds was [s]: M [SD] = 22,96 [15,90] and the decision time for simple decision making tasks in seconds was [s]: M [SD] = 13,61 [9,27].

Taking the total average decision making time into consideration it can be said that the difference between decision making times regarding task complexity is substantial. However, it is understandable because during complex decision making tasks the subjects had to process much more information than in simple tasks.

There was also a difference when the no rating condition and the rating condition were compared. If no ratings were presented to the subjects the decision time in seconds was [s]: M [SD] = 16,33 [11,19]. If ratings were presented to the subjects the decision time in seconds was [s]: M [SD] = 20,24 [14,35]. Considering the total average decision time, the difference between the no rating and the rating conditions is substantial. If ratings were presented the subjects spent 24% more time with the task than when no ratings were presented.

As per Hypothesis 4, there was a difference between the subjects’ decision time when bad ratings were presented and when good ratings were presented. When all the ratings were bad, the decision time in seconds was [s]: M [SD] = 19,5 [13,64]. When all the ratings were good, the decision in seconds was [s]: M [SD] = 20,98 [15,04]. If presented with good product ratings the subjects spent 7,6% more time with the decision making task than when bad product ratings were presented. A one-way ANOVA was also conducted with the time to first click as the dependent variable and quality of the product rating (good vs. bad) as the independent variable. The interaction between these variables did not reach a level of significance, F(1,398) = 1,062 (p= 0,303).

In order to explore the within-subjects variance a two-way repeated measures ANOVA was conducted with the time to first click as the dependent variable and task complexity and the quality of the product rating as the independent variables. The interaction of these variables reached a significant level, F(1,25) = 20,758 (p = 0,000). The combined effect of task complexity and quality of ratings had a significant impact on the time to first click. When the independent variables were explored individually, the interaction between time to first click and task complexity formed a significant main effect, F(1,25) =
7,071 (p = 0.014) and so did the interaction between time to first click and rating quality, F(1,25) = 27.048 (p = 0.000).

![Profile Plot, Time to First Click](image)

**FIGURE 4: Profile Plot, Time to First Click**

From the two-way repeated measure ANOVA, it can be conducted that when no rating was presented to the subjects and the task was complex, the subjects took a much longer time to process the task than if the task was simple and no rating was presented. Whereas, if ratings were presented to the subjects the decision times were a lot more similar.

**TABLE 4: Decision Times (Means and Standard Deviations)**

<table>
<thead>
<tr>
<th></th>
<th>n</th>
<th>M (s)</th>
<th>SD</th>
<th>Count</th>
</tr>
</thead>
<tbody>
<tr>
<td>All</td>
<td>25</td>
<td>18,284761</td>
<td>13,010236</td>
<td>800</td>
</tr>
<tr>
<td>Complex</td>
<td>25</td>
<td>22,95543</td>
<td>15,904031</td>
<td>400</td>
</tr>
<tr>
<td>Simple</td>
<td>25</td>
<td>13,614092</td>
<td>9.2756163</td>
<td>400</td>
</tr>
<tr>
<td>No Rating</td>
<td>25</td>
<td>16,326182</td>
<td>11.191172</td>
<td>400</td>
</tr>
<tr>
<td>Rating</td>
<td>25</td>
<td>20,243339</td>
<td>14.353505</td>
<td>400</td>
</tr>
<tr>
<td>Bad Rating</td>
<td>25</td>
<td>19,503625</td>
<td>13.635378</td>
<td>200</td>
</tr>
<tr>
<td>Good Rating</td>
<td>25</td>
<td>20,983044</td>
<td>15.035229</td>
<td>200</td>
</tr>
</tbody>
</table>
5.2.5 Time to First Fixation and Path Dependence

For the purpose of the time to first fixation metric the product ratings were treated as one common AOI rather than having one product rating represent one AOI. As per Hypothesis 1 and 2, it is of interest how fast the subjects perceive and start processing the product ratings but it was not of interest which one of the product ratings they would perceive first.

If a subject had not fixated on the AOI (product ratings) at all, the subject’s score for this measure was not included. For the purpose of this metric it was not of interest if the subject avoided – consciously or unconsciously – fixating on the product ratings and the value of 0 would have pulled the average of the time to first fixation down needlessly.

For all tasks where the time to first fixation was applicable (376 instances) the mean time in milliseconds was [ms]: M [SD] = 135,20 [60,07]. In complex decision making tasks where the time to first fixation was applicable (188 instances) the mean time in milliseconds was [ms]: M [SD] = 136,60 [64,10]. In simple decision making tasks where the time to first fixation was applicable (188 instances) the mean time in milliseconds was [ms]: M [SD] = 133,80 [55,93].

In order to explore the within-subjects variance a two-way repeated measures ANOVA was conducted with the time to first fixation as the dependent variable and task complexity and the quality of the product rating as the independent variables. The interaction of these variables did not reach a significant level, F(1,25) = 0,012 (p = 0,913). The combined effect of task complexity and quality of ratings does not have a significant impact on the time to first fixation on the AOI. When the independent variables were explored individually, the interaction between time to first fixation and task complexity did not form a significant main effect, F(1,25) = 0,303 (p = 0,587). However the interaction between time to first fixation and rating quality reached a significant main effect, F(1,25) = 6,250 (p = 0,020).

The comparison of mean times to first fixation indicate that, as opposed to Hypotheses 1 and 2, the subjects actually perceived the product ratings faster during simple decision making tasks. This could be due to the fact as the path dependence of the subjects shows, in most tasks the product information was assessed first and the product rating was assessed later. It is possible that the subjects used the product ratings more as a confirmatory tool rather than an actual factor for the decision during complex tasks, reflecting as a longer time to first fixation. In contrast, during simple decision making tasks the subjects had a lot less information visible to them, so it is possible that due to this they perceived the product ratings faster. It should also be noted that according to the subjects’ own assessments of their behavior during the tasks, many of the subjects were quite task-oriented as opposed to being oriented to leisurely shopping.

Hypothesis 3 postulated that in complex decision making tasks the subjects will perceive the good product ratings faster than the bad product
ratings. In complex decision making tasks where the time to first fixation was applicable (94 instances) and the rating was bad, the mean time in milliseconds was [ms]: M [SD] = 126,95 [59,05]. In complex decision making tasks where the time to first fixation was applicable (94 instances) and the rating was good, the mean time in milliseconds was [ms]: M [SD] = 146,24 [69,08].

Unexpectedly the subjects perceived the bad product ratings faster than the good product ratings in complex decision making tasks. This could indicate that when more information on the products is provided the more crucial the quality of the rating becomes. It is possible that the subjects did not expect to find bad ratings on products with a lot of information and this created a conflict, ultimately resulting in bad ratings being perceived faster.

To validate these results path dependence was used as a reflection of the subjects’ judgment and choice behavior. Venkatraman et al. (2014) posit that the subject’s use of a heuristic choice strategy with the aim of maximizing utility and minimizing the cognitive load will show up as path dependence. When comparing the personal assessment of each subject on the decision making strategy they used to the videos showing the scan path, it is possible to see scan patterns. The subjects were consistent in first comparing the product information between all models first, then moving on to comparing two of the models they considered to be the best and then to finalizing the decision making process through a comparison of product ratings. The images of the smartphones and the extraneous images on the fictional web store were not fixated on consistently, if at all. The subjects’ path dependence clearly indicates that they were trying to minimize their cognitive load through the use of a heuristic choice strategy.

### TABLE 5: The Main and Combined Effects of Task Complexity and Rating Quality on Eye-Tracking Parameters

<table>
<thead>
<tr>
<th>PARAMETER</th>
<th>Mean Fixation Duration</th>
<th>Fixation Density</th>
<th>Dwell Time</th>
<th>Dwell Time Percentage</th>
<th>Time to First Click</th>
<th>Time to First Fixation</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Two-way Repeated Measures ANOVA</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Task Complexity</td>
<td>F(1,25) = 0.384, p = 0.541</td>
<td>F(1,25) = 0.367, p = 0.550</td>
<td>F(1,25) = 0.009, p = 0.926</td>
<td>F(1,25) = 58,112, p = 0.000</td>
<td>F(1,25) = 7.071, p = 0.014</td>
<td>F(1,25) = 0.303, p = 0.587</td>
</tr>
<tr>
<td>Rating Quality</td>
<td>F(1,25) = 0.207, p = 0.653</td>
<td>F(1,25) = 0.018, p = 0.894</td>
<td>F(1,25) = 0.026, p = 0.874</td>
<td>F(1,25) = 0.001, p = 0.978</td>
<td>F(1,25) = 27.048, p = 0.000</td>
<td>F(1,25) = 6.250, p = 0.020</td>
</tr>
<tr>
<td>Combined Effect</td>
<td>F(1,25) = 0.489, p = 0.491</td>
<td>F(1,25) = 0.018, p = 0.894</td>
<td>F(1,25) = 0.127, p = 0.725</td>
<td>F(1,25) = 0.030, p = 0.864</td>
<td>F(1,25) = 20.758, p = 0.000</td>
<td>F(1,25) = 0.012, p = 0.913</td>
</tr>
</tbody>
</table>
6 DISCUSSION

The final chapter of this study is a discussion between the empirical findings of this study and the results of earlier research. The research questions set at the beginning of this study will be answered. The theoretical contributions of this study will be given and managerial implications will be covered. In this chapter the research done for this study will be evaluated in accordance with its limitations and in relation with previous studies. Finally, directions and proposals for future research will be given.

6.1 Theoretical Contributions

This study was conducted to answer the need for more research on the consumers’ path to purchase in an online context as well as on the cognitive influence and relevancy of different website design factors on consumer decision making. Thus, the following two primary research questions were set in the beginning of this study:

1. Do consumers rely cognitively on product ratings when making a purchase decision online?

2. Is the consumer purchase decision process affected after cognitively processing a product rating?

In addition to delve deeper into the first research question, two secondary research questions were applied:

1. Will purchase decision making complexity influence the need for affiliation among consumers when comparing products online?

2. Will purchase decision making complexity influence consumers’ use of cognitive heuristics when comparing products online?

This study tested four hypotheses in order to answer these research questions and gain insight on the effect of online word-of-mouth on consumer purchase decision making behavior.

Earlier studies (e.g. Just and Carpenter 1976, Orquin and Loose 2013) posit that increases in the working memory load (complex, cognitively
demanding tasks) should linearly increase the number of fixations as well as create more intentional re-fixations as well as that consumers generally try to minimize cognitive efforts when making decisions (Salant 2011; Wang et al. 2014). In the case of this study the amount of cognitive processing (fixation density, fixation duration and dwell time) and the speed of shifting attention to a social cue were used as measures to observe these phenomena.

Unexpectedly it was found that during a complex decision making task the subjects will perceive the product rating actually slower than a simple task as well as cognitively processed the product ratings more than during a simple decision making task, even when objective, non-social means of evaluation were available (as opposed to Hypothesis 1 and 2). Out of the parameters used to test these hypotheses only the results from dwell time were statistically significant.

Based on these results it would seem that consumers are very prone to using social heuristics even when the purchase decision at hand is seemingly quite simple. If given more information on products the consumers took a longer time to shift their attention from the product information to the product rating, possibly indicating that their priority was first to assess the objective information. In addition, the fact that consumers tend to compare their opinions socially when faced with uncertainty (e.g. Festinger 1954; Taylor et al. 1990; Gibbons and Buunk 1999), came into play in simple tasks. When interpreting the verbalization of the subjects’ thought processes, in simple tasks the subjects were filling information gaps, i.e. uncertainty about the missing product information, by using the product ratings.

Many previous studies (e.g. Park et al. 2007; Gupta and Harris 2010; Wolf and Muhanna 2011; Flanagin et al. 2014) have postulated that consumers associate high product ratings with a higher level of trust and product quality and increase the likelihood of purchase. As per Hypotheses 3 and 4 during a complex decision making task the subjects perceived the good product rating faster and cognitively processed it more than a bad product rating and the presentation of a good product rating influenced the subject’s cognitive and decision making process more than the presentation of a bad product rating, regardless of task complexity. When exploring the decision time per subject, statistically significant results show that if presented with good product ratings the subjects spent more time with the decision making task than when bad product ratings were presented.

There was also a significant difference when the no rating condition and the rating condition were compared. If ratings were presented the subjects spent more time with the task than when no ratings were presented. Gupta and Harris (2010) have also found out in their earlier research that e-WOM increases the time a consumer considers the recommended product. It could also be conducted that when no rating was presented to the subjects and the task was complex, the subjects took a much longer time to process the task than if the task was simple and no rating was presented. Whereas, if ratings were presented to the subjects the decision times were a lot more similar. In other words, ratings closed the complexity gap between the two different types of tasks. These findings are similar to for example Cyr et al. (2007) and Fang (2012)
who found out that by including elements that implement the use of social influence, online reviews, or online interactivity, consumers are able to find extra cues to help them in the purchase decision making process.

6.2 Managerial Implications

There are several managerial implications that can be drawn from this research. Firstly, as managers are designing the online customer experience of a web store and building a path to purchase, electronic word-of-mouth serves a large purpose of beneficial social comparison for the consumer. Not only do consumers associate strong e-WOM with a higher level of trust and product quality (Park et al. 2007; Wolf and Muhanna 2011; Flanagin et al. 2014), but they also spend more time exploring the products at hand when e-WOM is presented (Gupta and Harris 2010). If e-WOM is not presented to the consumer, for example in the form of product ratings, they will spend less time assessing the products, endangering purchase decision quality. Presenting e-WOM to consumers is especially important when the purchase decision making situation is cognitively demanding, for example in the case of technical products with multiple attributes to assess. In these situations, consumers are prone to attend to strong product ratings.

Consumers make many decisions based on social heuristics and use them to simplify the purchase decision making process. (Wästlund et al. 2015). By including social influence, online reviews, or online interactivity, consumers are able to find additional cues to help them in the purchase decision making process. (Cyr et al. 2007; and Fang 2012). Therefore, it is vital for managers to assist the consumers in this process by providing relevant cues that trigger the use of heuristics. Even if the purchase decision should seemingly be simple to make, consumers still look to cues such as e-WOM in the form product ratings to assist them in the decision making process.

Regardless of purchase decision making complexity, product ratings – good or bad – provide additional help to consumers. Consumers shift their attention to these ratings in a matter of milliseconds when presented on a web store; sometimes even disregarding objective means of decision making. Hence, especially in the case of strong product ratings, it seems only logical to include them on the web store in a setting where attention is easily shifted towards them.
6.3 Evaluation of the Research

According to Metsämuuronen (2005, 12–14) and Krauth (2000, 21) there are several threats for validity when conducting experimental research. These can be divided into 1) statistical conclusion validity, 2) internal validity, 3) construct validity of putative causes and effects, and 4) external validity.

Several measures were taken to mitigate the threats to statistical conclusion validity such as the independence of measurements, preventing random disturbances (e.g. extraneous noise from outside the experimental setting) as well as randomization in terms of product attributes and product ratings. As statistical power is derived from the sample size it must be stated that for the purpose of this study the sample size was adequate but in order to achieve statistically more powerful results, the sample size should be larger. The reliability of the experiment was improved through the standardization of the treatments per each subject. As the aim of an experimental study is to control the variables, and through this gain validity, the independent variable should be the cause of the effect or result. In the case of this study the independent variable can be seen as a cause of for the desired effect. (Metsämuuronen 2005, 12-14; Krauth 2000, 22-25.)

Internal validity is mainly concerned with causality, i.e. whether factor a causes factor b. In the case of this study in terms of internal validity, the historical, maturation and regression effects that threaten the validity of research do not affect the results of the study, even though a control group was not used for this study. (Metsämuuronen 2005, 12-14; Krauth 2000, 25-27; Yin 2014, 46.)

Construct validity is concerned with the exactness of the constructs and definitions used in the study. As per construct validity the inexactness of the definition of constructs in a prominent threat. However, in the case of this study measures such as careful and thorough definitions of constructs based on earlier research were taken to mitigate the effects of this threat. (Krauth 2000, 29.)

In terms of external validity, or in other words the generalization of the results, the subjects who participated in this study were from a not well defined subpopulation, so it is questionable whether the results of this study would hold for a subpopulation for which no previous studies have been conducted for. (Krauth 2000, 33; Yin 2014, 46.)

Based on the evaluation of this research and the measures taken to mitigate the effects of various threats on validity and reliability, it can be concluded that the internal and external validity and reliability of the results of this study are adequate.
6.4 Limitations of the Research

Several limitations need to be taken into account when interpreting the results of this research. First of all, it needs to be acknowledged that even though this study was conducted based on an objective research method (eye-tracking), the results were collected during a simulated experiment. Therefore, it cannot be said that the research would have fully reflected a real-life setting where a consumer is browsing through a web store. The web store the subjects used was fictional, and even though it was made to be a realistic as possible it still lacked certain elements commonly found in a web store. In addition, the presentation of the web store had to be optimized for the eye-tracker, which caused the web store to lose interactive elements. The eye-tracking equipment also forced the subjects to sit with their forehead pressed against the equipment and their chin resting on a chin rest, which is not a realistic way of behaving when browsing a web store.

Secondly, even though eye-tracking has been proved to be a method of collecting valid data and to be able to collect objective information from subjects’ gaze behavior, the data still needs to be interpreted by the researcher. Therefore, it cannot be stated that the data collected would be conclusive. (Waechter et al. 2015.)

Thirdly, although appropriate measures were taken to strip the web store and the experimental setting from any additional and extraneous stimuli, some had to be left in place to create a sense of realism to the experiment. It is possible that these additional stimuli, such as the pictures of smart phones or the general design of the web store affected the gaze behavior of the subjects and furthermore the results. Also in accordance with the descriptions the subjects gave verbally about their purchase decision making strategies many of them had evaluated all information from top to bottom. The product ratings were positioned after the product information and sat at the bottom of the web page. This may have affected the results.

Also, the subjects participated in the experiment at different times, so it is possible that they under different levels of fatigue and stress. Also their motivation to fully participate in the experiment as realistically as possible might have differed, affecting the results.

Finally, it needs to be acknowledged that that sample consisted mostly of young people (aged 18-25) and every subject was Finnish. All subjects were also familiar with using web stores and smartphones, which may not always be the case among consumers. Therefore, the sample is not fully random in terms of all background factors, which may have affected the results. (Metsämuuronen 2005, 14.) The study was conducted in the context of smart phones – being a relevant product to a sample of young people – so there is no guarantee that the results would be exactly the same in another product category or context.
6.5 Directions for Future Research

Understanding customers and the customer experience in a world shaped by digital technology will be paramount in the future (MSI Research Priorities 2014-2016). Even more so, there exists a need for more research on the consumer’s cognitive processes when making decisions, the influence of website design factors and objects towards consumer decision, providing consumers with helpful information online as well as the social presence in online environments (De Vries and Pruyn 2007; Chae and Lee 2013; Shi et al. 2013; Roth et al. 2013; Ben et al. 2015).

As was mentioned in the limitations to this research, the experimental setting created some limitations for this study. In terms of future research, a setting where subjects could feel less task-oriented and more shopping oriented might uncover interesting results. In future research on this topic the fictional web store could be made to feel even more realistic to create a setting, which matches a real-life online shopping experience. In addition, for future research acquiring a larger sample size to validate these results further would be important.

This research showed only a variety of smartphones to the subjects to choose from. Future research should investigate whether similar results could be found in other product categories, and possibly in product categories with less technical attributes such as clothing. The research could also be broadened to different age and nationality groups in terms of the subjects. Future research could also support the eye-tracking data with a wider supporting behavioral data.

The importance of price has to be acknowledged from the written answers on the strategies the subjects used to make their purchase decisions. Even though the prices of the products were made as similar as realistically possible, it would be of interest to conduct a study where the effect of price is fully neutralized.

It would also be of academic interest to explore the effects of electronic word-of-mouth on product choice more thoroughly. This study was more focused on the shifts of attention towards the ratings but the actual effect of the ratings on consumer behavior remains to be a curious topic in the field of eye-tracking research.
REFERENCES


Microsoft. 2013. The consumer decision journey: Understanding consumer decision-making along the retail path to purchase. Microsoft Advertising Consumer Insights.

Miller, G. 1956. The magical number seven, plus or minus two: some limits on our capacity for processing information. The Psychological Review 63 (2), 81–97.


Russo, E. 1978. Eye fixations can save the world: A critical evaluation and a comparison between eye fixations and other information processing methodologies. Advances in Consumer Research 5, 561-570.


## APPENDIX 2: PRODUCT DESCRIPTIONS: SMART PHONES

<table>
<thead>
<tr>
<th>Product</th>
<th>Display</th>
<th>Camera</th>
<th>Memory</th>
<th>Network Connection</th>
<th>Price</th>
<th>Battery</th>
<th>Design</th>
</tr>
</thead>
<tbody>
<tr>
<td>Waterproof Smartphone Red</td>
<td>4.8” HD</td>
<td>16 MP</td>
<td>64 GB</td>
<td>4G/LTE</td>
<td>399.99</td>
<td>1200 mAh</td>
<td>Slimline</td>
</tr>
<tr>
<td>Waterproof Smartphone Silver</td>
<td>5.0” Full HD</td>
<td>12 MP</td>
<td>128 GB</td>
<td>5G</td>
<td>499.99</td>
<td>1800 mAh</td>
<td>Compact</td>
</tr>
<tr>
<td>Waterproof Smartphone Black</td>
<td>5.5” Super AMOLED</td>
<td>24 MP</td>
<td>256 GB</td>
<td>4G/LTE</td>
<td>599.99</td>
<td>2100 mAh</td>
<td>Sleek</td>
</tr>
<tr>
<td>Waterproof Smartphone Blue</td>
<td>6.0” QHD+</td>
<td>32 MP</td>
<td>512 GB</td>
<td>5G</td>
<td>799.99</td>
<td>3000 mAh</td>
<td>Premium</td>
</tr>
<tr>
<td>Waterproof Smartphone Green</td>
<td>5.2” 4K OLED</td>
<td>36 MP</td>
<td>1 TB</td>
<td>4G/LTE</td>
<td>899.99</td>
<td>3500 mAh</td>
<td>Luxurious</td>
</tr>
</tbody>
</table>

*Product specifications subject to change without notice.*
APPENDIX 3: BACKGROUND INFORMATION SHEET FOR SUBJECTS IN ENGLISH

University of Jyväskylä  
School of Business and Economics  
Background sheet

<table>
<thead>
<tr>
<th>NAME:</th>
<th>GENDER: M / F</th>
</tr>
</thead>
<tbody>
<tr>
<td>AGE:</td>
<td>OPTICAL PROBLEMS: YES / NO</td>
</tr>
<tr>
<td></td>
<td>If yes, please describe briefly:</td>
</tr>
</tbody>
</table>

PREVIOUS WEB STORE EXPERIENCE  
(circle the number that corresponds to you):

1 – I use web stores daily  
2 – I use web stores weekly  
3 – I use web stores approximately once a month  
4 – I use web stores a few times per year  
5 – I use web stores less than once a year  
6 – I have never used a web store

PREVIOUS SMARTPHONE EXPERIENCE  
(circle the number that corresponds to you):

1 – I use smartphones daily  
2 – I use smartphones weekly  
3 – I use smartphones approximately once a month  
4 – I use smartphones a few times per year  
5 – I use smartphones less than once a year  
6 – I have never used a smartphone

SMARTPHONE TECHNICAL ATTRIBUTES  
(circle the number that corresponds to you):

1 – I know the technical attributes of smartphones very well  
2 – I know the technical attributes of smartphones quite well  
3 – I don’t really know the technical attributes of a smartphone  
4 – I don’t know the technical attributes of smartphones at all

SIGNATURE