

Master Thesis Marketing

Do we interpret the online review differently?

What are the effects of different review components?

Name	Charlotte Boons
Student number	454541
Supervisor	Prof. Dr. M.G. de Jong
Second assessor	Prof. Dr. W.J.M.I Verbeke
Faculty	Erasmus School of Economics
Date	23 April 2018



ERASMUS UNIVERSITEIT ROTTERDAM

Abstract

Due to the introduction of the internet, WOM occurs no longer solely offline but online as well. As result of the increased accessibility of being online present, the usage of online reviews is tremendous. Reading and sharing online reviews became a common activity. It has been found WOM could influence pre-purchase decisions (Prunawirawan et al., 2012). Especially, during online as well as offline purchase decisions, online reviews play a key role. It has been found that the review function of delivering trustful information is most necessary in case of purchasing experience goods. This study is based on the movie industry which belongs to the experience goods on account of the difficult way of predicting the movie experience. The influence of online reviews in the industry of movies became more important, due to the increased accessibility of watching movies and of collecting and sharing movie reviews as consequence of the birth of many movie review platforms.

The aim of the current study was to examine the different effects of several review components in relation to the perceived review credibility and the willingness to watch a movie. Derived from realistic components of movie reviews, the following components were examined: star rating, text valence, and sources. The outcomes of this study are based on survey results from 520 respondents. The conducted survey contained 13 conditions.

The main finding of this study is that depending on the combination of review components, star rating, review text and review source, an online movie review has an impact on the willingness to watch a movie. The results suggest that star rating and review text play different roles depending on whether a review consists of a star rating as well as which source had written the review. The star rating seems to be dominant in case the online review is written by a user. The opposite effect occurs in case the online review is written by an expert, whereby the review text has the most impact. The main findings related to perceived review credibility were not sufficient enough to get valuable insights.

The main benefits of the findings are the many insights on the domain of online reviews. The findings of this study provides the opportunity to understand what happens in the mind of the potential consumers during the interpretation of online reviews. Keywords: Online reviews, review components, decision making, online review interpretation, review source, Electronic word-of-mouth (eWOM).

Table of contents

Abstract.....	2
1. Introduction.....	5
1.1 Background information.....	5
1.2 Problem statement and research question.....	7
1.3 Research questions	7
1.4 Academic relevance.....	7
1.5 Managerial relevance.....	8
1.6 Structure of the thesis	10
2. Literature.....	11
2.1 Online decision making.....	13
2.2 Willingness to watch a movie.....	15
2.3 Credibility.....	16
2.4 Review valence.....	17
2.5 Star rating and written reviews.....	19
2.6 Review source.....	22
2.7 Conceptual Framework.....	25
3. Research Methodology	27
3.1 Research design	27
3.2 Sample	27
3.3 Experiment overview.....	28
3.4 Measurements.....	30
3.5 Manipulation check	31
3.6 Statistical methodology	31
Chapter 4 Results	33
4.1 Assumptions ANOVA.....	33
4.2 Analysis of variance	33
4.3 Main effects ‘Willingness to watch’.....	34
4.3.1 Review use.....	34
4.3.2 Rating and written review	35
4.3.4 Source.....	36

4.4 Main effects ‘Credibility’	39
4.4.1 Valence.....	39
4.4.2 Star rating.....	40
4.4.3 Congruousness and Incongruousness.....	41
4.4.4 Source.....	41
4.3.4 Halo-effect.....	42
5. Discussion.....	45
5.1 Relation to existing findings.....	45
5.2 Academic implications	46
5.3. Practical implications	47
5.4 Limitations and Future Research.....	48
6. Conclusion	51
7. References.....	52
Appendix 1 – Pre-test.....	61
Appendix 2 – Pre-test Outcomes	65
Appendix 3 - Final Survey.....	66
Appendix 4 - Statistical Outcomes	72
Matrix: Willingness to watch a movie.....	74
Matrix: Review credibility.....	76

1. Introduction

1.1 Background information

Imagine you and your friends are planning to watch a movie together. Before you all decide which movie you want to watch, you and your friends start orienting on the possibilities. You read a lot of different reviews and, based on that, you make a selection of the best options. During the orientation part, you and your friends first see the star-rating score and the review source. After that, you all read the written part of the reviews. Finally, you make a decision which movie you will watch. Now that you have identified this decision-making process, it might be interesting to wonder the following, what role do the reviews play in the decision-making process and how do a rating system and difference in review sources influence the final evaluation? For example, if you and your friends see a review that contains a relatively high rate, you may interpret the written part differently because you may read the same written part with different eyes. This also applies in a case where same reviews are written by different sources.

Consumers purchase decisions consist of many stages, and marketing plays a crucial role in some of these (Edelman, 2010; Hoeffler & Ariely, 1999; Philips et al., 1995). Before consumers select the online reviews, their attitudes towards the alternative product options are essential and are influenced by marketing efforts (Edelman, 2010; Priester, Nayakankuppam, Fleming, & Godek, 2004). During the final stage, consumers reduce the set of options to a small choice set or to a single product or service choice. While this consideration occurs, the key factors such as product experiences, might lay outside the influence of the marketers (Hoeffler & Ariely, 1999).

Nowadays, the use of the internet to purchase products and services is growing strongly and consequently the use of online reviews is steadily increasing. This can be explained by two factors. First, the growing availability of commercial websites and, second, the acceptance of online transactions by consumers (Hong et al., 2004). The main disadvantages of online shopping are that consumers do not have the opportunity to touch, see, or try products or services. As a consequence, consumers are increasingly dependent on other information. Because of this need, online sellers created the possibility for consumers to share their product and service experience with others on various websites (Avery et al, 1999).

Nowadays, consumers have the opportunity to collect seller created product information as well as buyer created product information. The latter is displayed as an online review, which is based on the consumers' personal usage experience, product evaluations and opinions.

As mentioned above, consumers search for external information such as online reviews before they purchase. Especially when the quality of goods is hard to predict the need for consumers for external information will be higher (Hao et al., 2010; Nedungadi, P. 1990). Therefore, purchase decisions of experience goods such as movies are strongly influenced by online reviews (Lynch & Ariely, 2008; Hoch & Ha, 1986). In the movie industry, online movie reviews mostly consist of two different parts: a rating score and a written contribution. The rating is often based on a five-star system and provides consumers with a quick overview of the written content. It can also help consumers to decide which written review they will read. The written part of the online review is where consumers explain their product evaluation and opinion based on their product usage experience. In general, the two separate

parts should be aligned. Obviously, consumers are different persons with individual opinions and distinct priorities, which may result in different outcomes between the movie rating scores and the written review. This potential discrepancy is due to the fact that the rating part may have an impact on the way consumers will read the written part. This alleged discrepancy may affect the final perceived review credibility and the willingness to choose a product.

Because of the opportunity of using the internet for sharing product information, not only user experiences and opinions can be shared but also those of experts. These experts refer to people with special expertise regarding the products or services. Eliashberg and Shugan (1997) were the first to propose and test two different roles of experts: the influencer and the predictor. An influencer, or opinion leader, is a person who is regarded by a group or by other people as having key expertise or knowledge on a particular subject. The size of the impact that the expert reviews have on the decision-making process could depend on the perceived credibility of the total review. The credibility of the review could be influenced by the two different sources: user and expert. For example, when you and your friends are reading a movie review from an expert source, you could think this expert review has more credibility than a movie review written by users because an expert has the responsibility to protect his title as an expert and has more product knowledge than you have.

On the other hand, you could also think that you are more connected with other users because you have the same position. Moreover, users do not have extra interests to manipulate you as a consumer, thus increasing your belief in their evaluations and opinions.

1.2 Problem statement and research question

Is the willingness to watch a movie and the perceived credibility of movie reviews influenced by different online review components?

1.3 Research questions

1) Is the willingness to watch a movie and the perceived credibility of movie reviews influenced by difference in online review valence?

2) Is the willingness to watch a movie and the perceived credibility of movie reviews influenced by difference in star ratings?

3) Is the willingness to watch a movie and the perceived credibility of movie reviews influenced by difference in online review sources?

1.4 Academic relevance

The research findings related to online reviews are mostly about motivations and social dynamics between users and contributors of online review websites (Hennig-Thurau, Gwinner, Walsh, & Gremler, 2004). In addition, recent studies indicate the helpfulness of online reviews (Zhou & Guo, 2017; Agnihotri & Bhattacharya, 2016; Schindler & Bickart, 2012). But these studies do not consider the impact of the online reviews on consumer decision making and to what extent online reviews affect the attitudes of consumers and decisions (Chevalier & Mayzlin, 2006; Sen & Lerman, 2007). Furthermore, existing research proved the effect of online reviews on product sales (Hu, Liu and Zhang, 2008). This research took the role of reviewer characteristics as well as temporal effects into account. However, this research did not make the nuance in the various review components and examine the effects of these components on product sales.

Many studies examined the different effects of online trust and especially with respect to online purchase behaviour (Banes & Chen, 2007; Gefen, Karahanna & Straub, 2003; Jarvenpaa, Tractinsky & Vitale, 2000; Garbarino & Johnson, 1999; Hoffman, Novak & Peralta, 1999). But none of these studies examined the different effects on perceived review credibility of various components of online reviews.

This research will contribute new findings on this relation between online review components and the willingness to choose a product and the perceived credibility of a review. The results of this study could lead to be a reason and give a direction for future research. The contribution of this research will not only be valuable from a theoretical perspective but also from a managerial perspective.

1.5 Managerial relevance

The abovementioned situation of people searching for online movie reviews, could occur more frequently due to the consistent growth of the movie industry. The worldwide box office revenue for the movie industry will grow from 9.767 billion U.S. dollars in 2007 to 10.836 billion in 2017 (Nash Information Services, 2017). The global online television and movie industry will have a revenue amount of 2.1 billion U.S. dollars in 2018 and will have an increasing role in the total movie industry in the future (Statista, 2017). Because of the increasing importance of the online television and movie industry, new websites such as 'Netflix' and 'MetaCritic' have emerged. As a result, it has become necessary for movie makers and cinemas to spend more money on advertising to reach consumers for the promotion of their newly-released movies (Statista, 2017). At the same time, online recommendations become indispensable for the upcoming online movie industry. In other words, in the online as well as the offline movie industry, it is of high importance to determine the influence factors on consumer movie preference. The online movie reviews therefore play an increasingly important role in this industry.

In addition, the use of online movie reviews will not only grow as result of the described developments in the global movie market but the general use of online evaluations and opinions also play a strong role. Sharing and searching online reviews represent one of the fastest growing online phenomena (Armstrong & Hagel 1996). The increased need for online reviews is related to the need for credible information. Nowadays, people are aware of the trust-issue of

online information, which refer to the lack of means to check whether the information is true. Several studies indicated that the lack of trust is one of the greatest barriers during internet transactions (Hoffman et al., 1999). Furthermore, the extent of trust perceived by potential customers could affect the purchase intention (Gefen et al., 2003; Jarvenpaa et al., 2000), repeat purchases (Garbarino & Johnson, 1999) and the loyalty of customers (Gefen, 2002). This makes it clear that the extent of trust in product or service information before purchase has a key role in the purchase decision. With this, it is relevant for any type of supplier to know how to create a higher perceived trust. Current research shows that a great part of online reviews is written by firms themselves by writing 'user reviews' and are not trustworthy (Jindal & Liu, 2008). In addition, the reliability of expert reviews is questionable, because firms can easily manipulate these online evaluations of experts by sponsoring these experts. Obviously, firms do not tend to pay experts who could give negative online evaluations. This emphasizes the need for knowing how online reviews are interpreted and what and to what extent it has an impact.

The importance of the online reviews role during purchase decisions is increased. Possibly, the most important reason for this is according to Chen and Xie (2008), that online user reviews can be exploited as a new element in the marketing communication mix. The authors describe such user reviews as 'sales assistants'. Consumers can use these free 'sales assistants' to identify their needs because these online reviews can have a big impact on purchase decisions. For sales managers, it would be necessary to gain huge control on this new marketing communication element to be able to use it optimally whereby a better market position and a sales boost are the consequences (Park et al., 2007).

For retailers, Lynch and Ariely (2000) research results indicate who carrying differentiated goods several incentives to make information environments maximal transparent as possible. Anderson, Palen and Vieweg (2010), conclude in their study about the everyday analysts, that the everyday analyst (the information seekers or potential consumers) have the need to receive support while they consider the information content and credibility, and finally making decisions across multiple sources of information. The everyday analyst stays in control, meaning that making decisions is not the job of the information given tools and services but still of the everyday analyst. However, the information given tools and service have to improve the way how the everyday analyst makes decisions to make it as confidently and quickly as possible (Anderson et al, 2010). Lynch and Ariely (2000) and Anderson et al, (2010),

underlined the need for making ‘online information searching and selecting’ easier for potential customers. The findings of this research will contribute to answering this need.

After reading this research, marketers, review platform suppliers, consumers, experts and marketing academics will better know how online reviews influence the willingness of a product and the credibility of online reviews by the concerned online review components. Especially, marketers, online review platform suppliers and experts, can use these findings to better achieve their goals. For instance, marketers will have the opportunity to interpret and act better the existing online reviews as an unaware consumer, which could explain the current online consumer behaviour. Review platform suppliers should know how to customize their platform to provide it more accessible and user-friendly for visitors by eliminating and focussing on some review components, which may lead to a higher platform traffic. Furthermore, consumers are benefit from the more use friendly customizations of the review platforms as well. The threshold to using online review will be reduced. On the other hand, consumers will have a better know how online reviews have an impact on their one behaviour what may lead that they think they have more control over their behaviour. Finally, after reading this study, review writers will know which review components will and will not support their goals. For instance, user writers will have a focus on other review components compared to expert writers.

1.6 Structure of the thesis

In order to answer the research question, this study consists of three parts. In the first part, a mapped view of the current relevant literature is provided. This literature is the basis of the formulated hypotheses, which are being tested in the second part of this study. The data will be obtained from available online experimental questionnaires. By using statistical analyses, the answers will get a more trustful meaning. In the last part of this thesis, research results will be discussed and will be followed by a general conclusion. This will address the outcome of the study and provide an answer to the defined research question, the state of managerial implications and academic implications and the suggestions for future research.

2. Literature

The following chapter is the foundation of this research. In parts 2.1 to 2.6 of this chapter, the variables and the hypotheses of the current research are derived from prior research findings. After this, the total overview of the variables and their expected relation will be mapped in the conceptual framework that is provided in paragraph 2.7.

The expected effects based on the existing research findings are based on the assumption that people reading one review and interpret the complete review. To understand the content of the following formulated hypotheses, the following table of definitions and abbreviations must be taken into account.

Table 1.
REVIEW CONDITIONS

Definition	Abbreviation	Movie Review	Star Rating	Text Valence	Source
Willingness to watch a movie	W2W				
Credibility of review	CR				
Positive review	P	Yes	No or positive	Positive	Expert or User
Negative review	N	Yes	No or Negative	Negative	Expert or User
Expert review	E	Yes	No, Positive or Negative	Positive or Negative	Expert

User review	U	Yes	No, Positive or Negative	Positive or Negative	User
Baseline	B	No	No	No	No
Congruent	PPE	Yes	Positive	Positive	Expert
	PPU	Yes	Positive	Positive	User
	NNE	Yes	Negative	Negative	Expert
	NNU	Yes	Negative	Negative	User
Incongruent	PNE	Yes	Positive	Negative	Expert
	PNU	Yes	Positive	Negative	User
	NPE	Yes	Negative	Positive	Expert
	NPU	Yes	Negative	Positive	User
Only Text	_PE	Yes	No	Positive	Expert
	_PU	Yes	No	Positive	User
	_NE	Yes	No	Negative	Expert
	_NU	Yes	No	Negative	User

2.1 Online decision making

The traditional consumer decision-making process can be divided into four steps. First, consumers start with finding and selecting potential brands. This is followed by the identification of the relevant attributes of each selected brand. After the identification of the several different attributes of the various brands, the evaluation will occur as step three. The traditional process will finish by reducing the number of potential brands, selecting the final brand and making the purchase complete (Philips et al., 1995).

However, due to the growth of the digital economy, the traditional consumer making process is changed. Nowadays, consumers are able to read online feedback and ratings, to use gestural controls and evaluative interactive videos or 360-degrees views that demonstrate the product (Ramaswamy, 2013). The decision-making strategy 'Satisficing' explains the typical behaviour of decision makers (Simon, 2013). According to Simon (2013), decision-makers cannot choose an optimal solution. Instead, people settle for the first adequate solution they can find. Often this is a necessary approach in cases in which alternatives are presented one by one or for when there are multiple different choices. When people are searching for information online, they will be overloaded by information about alternatives and multiple choices. For instance, when people are searching online for information to base their decision on, satisficing will not only be present during the movie choice, but also during the decision whether they take a movie review into account. Therefore, the movie choice, as well as the review choice, do not have to be the optimal solution for consumers. Instead it has to meet their minimum requirements (Simon, 2013).

Then the question rises up: how can people find the best trustworthy information before they decide? Because of the accessibility of the internet, consumers may decide to search information by using internet platforms and to share their own evaluations of different products/services online (Avery et al., 1999). With this, word-of-mouth marketing (WOM), occurs not only offline but also online in the form of online reviews. Product evaluations that are available on a company's and external party websites are online reviews (Mudambi & Schuff, 2010). The relation between WOM and online reviews is based on the same direct way of information sharing between people. Because online reviews became one of the most important channels, it is generating more online WOM (Duan et al, 2008). It could influence purchase decisions in a negative or positive way (Welke, 2015). Hammond, East and Lomax (2008), concluded that the impact of both directions (positive/negative) are related to pre-WOM

purchase probability and the strength of WOM. Solicited WOM has more impact on decisions than non-solicited WOM. Online reviews are examples of solicited WOM because people actively search for this external information/advice.

Online solicited WOM often takes place during online purchase decisions. According to the findings of the study 'The effect of Word of Mouth on Sales: Online Book Reviews' (Chevalier & Mayzlin, 2006), WOM affects the online purchase behaviour of consumers. This is underlined by 'The Customer Decision Journey' (Edelman, 2010; Philips et al., 1995), which emphasizes the important role of online reviews during the customer decision journey. Edelman (2010, p. 3) suggests that 'the coolest banner ads, hottest viral video's and best search buys may win consideration for a brand, but if the product gets weak reviews, or is not even discussed online, it is unlikely to survive the winnowing process'. According to Court et al (2009), the approach of consumers searching and buying products or services is changing. This change is driven by the new technology and internet developments. The current consumers do not act passively, for instance waiting for advertising is coming. Instead, consumers are acting actively by searching for and reaching to online blogs, websites and online reviews to distinct what their options are (Court et al, 2009). Online reviews affect the purchase intention of consumers by strengthening or weakening the willingness to buy a product or service. This is underlined by the following recent findings: before consumers purchasing expensive products and or products with difficult predicted product experiences 66% of those are reading online reviews. Besides this, 88% of online shoppers said that they incorporate online reviews into their purchase decision making process (Dimensionalresearch, 2013; Wüthrich, 2015).

In order to understand why the information from others could be important during the online purchase decisions, research showed the need for additional information depending on how easy or difficult the certain decision is (Gino & Moore, 2007). People are more likely to follow the recommendations of others when the level of knowledge of decision makers is low. In case the level of knowledge is high, they are less likely to follow the recommendations made by others. This is caused by the differences in need for external input (Yaniv, 2004). Online reviews provide two different information types. First, online reviews give general product information. Secondly, online reviews act as a recommendation. Last mentioned is considered as most important information mainly because of the personalized content (Wang et al., 2012).

For some products it is more difficult to predict the attributes regarding how the use of these products will be experienced. These experiences refer to certain product characteristics

including the price-quality ratio (Nelson, 1970). Experience goods have the need to be tried in reality in advance of the purchase. Examples of experience products are video games, CDs and movies. For the last experience product, *i.e.* movies, consumers use multiple online reviews to predict the experience of it. Search goods have the opposite character, due to the experiences being easier to predict (Hao et al., 2010). It is also easier to find information on the quality of a product for search goods before purchase take place (Nelson, 1974; 1970).

Because of the complex prediction of product experience of experience goods, consumers tend to search for alternative information sources such as product tests and recommendations (Hoch & Ha, 1986). As a result, WOM becomes a more important driver of the consumer's demand for such products. The complexity of predicting the product experience before the purchase is decreasing as a result of the internet use which makes it possible to discover the quality of product experiences easier (Lynch & Ariely, 2000). Due to the rapid expansion of the internet, the WOM is no longer limited by geographical-social circle and timespan (Nelson, 1970). The use of the internet for conducting of WOM resulted in the birth of a new definition of WOM; eWOM, which stands for electronic word-of-mouth (Kietzmann & Canhoto, 2013). This all means that WOM is more needed but at the same time more available than ever. Duan et al. (2008) as well as Park and Lee (2008), indicated that experience goods are the most sensitive for online reviews. According to Huang et al. (2009), online reviews have a larger impact in the domain of experience goods compared to search goods. Research shows that consumers prefer a brief description on objective elements of the products and services. In addition, consumers prefer a brief description about highly sentimental components of the products and the services that are not recorded by the product and services description (Ghose & Ipeiroitis, 2006). It would therefore be obvious if consumers are more demanding, regarding information in reviews for experience products than for search products.

2.2 Willingness to watch a movie

Existing research suggests that purchase intention is a relevant indicator in the field of marketing (Morrison, 1979). Purchase intention is the likelihood of a consumer purchasing a product and service (Whitlark et al., 1991). By measuring the purchase intention, marketing managers are able to receive a clear view of the consumers' future plans (Morwitz, 2012; Ajzen, 1991). The existing theories suggest that intentions of individuals take part in the behaviour.

This indicates that it is possible to predict actual behaviour of individuals (Morwitz, 2012; Fishbein & Ajzen, 1975). These theories have been tested in the domain of marketing, resulting in the purchase intention being relevant in the marketing field (Morrison, 1979).

There are multiple websites that are offering different types of online reviews to help the consumer decide. At the same time, an expanding number of websites like 'Netflix', 'YouTube' and those of television stations, increase the accessibility to watch a movie by offering a wide range of movies for the relatively good price. Nowadays, consumers are not forced anymore to spend money for each movie they want to watch but they have the opportunity to use the convenient, monthly payment system. As a result, it is more relevant to measure 'willingness to watch' instead of 'purchase intention' or 'willingness to pay' in this industry.

2.3 Credibility

The increased use of the internet has been accompanied by the online trust issue. The internet is accessible for everyone that has led to a general suspicion, especially around online transactions. There is a positive impact of online trust on purchase intentions (Barnes & Chen, 2007). This is because people's purchase decisions are influenced by the extent in which consumers base their decisions on the online external information. Online trust mediates, at least in part, the relationship between behavioural intent, website and consumer characteristics (Yakov, et al Oct. 2005). Furthermore, online trust as mediation is strongest for websites which are oriented toward infrequently purchased, high-involvement items such as computers. Research shows order fulfilment and privacy are the most influential determinants of online trust of websites in which both involvement and information risks are high. For information-sensitive websites, navigation is strongest influential determinants. For websites with relatively high-involvement categories, brand strength is critical (Yakov, et al Oct. 2005).

In 2015, 66% of 30,000 global online respondents (scope of 60 countries) completely or somewhat trusted consumer opinions posted online, compared to 70% who completely or somewhat trusted branded websites (Nielsen, 2015). Although the mentioned percentages are not based on scientific research, these findings suggest that people are tending to trust branded websites and consumer's opinions posted online. However, other findings of this research suggest that the level of online trust will decrease in the future. This development already took occurs in the developed countries in 2015 (Nielsen, 2015). People become more aware of the

reliability risks of the dependence of others maliciousness and/or approachability to send no reliable information.

In case online reviews do not contain enough necessary information, they will be less likely to be perceived as reliable (Ratchford et al, 2001). Furthermore, previous research showed that credibility could play two different roles: 1) the influencer or 2) predictor during the decision making process. The level of the perceived credibility could be responsible for the effect size of the influences or prediction on the decision making (Chatterjee, 2001; Hu et al., 2008; Sen & Lerman, 2007; Xia & Bechwati, 2008). Based on the prior research findings, there is reason enough to take the credibility of online reviews into account. Consequently, the following questions arise; ‘which factors determine the extent of the perceived review credibility?’, ‘what is the order of impact of these factors on the perceived credibility?’ and ‘what are the effects of combinations of these factors on the perceived credibility?’. For answering these questions, it is necessary to focus on the review itself, similar to the study about website trust by Yakov, et al (2005). First, the perceived review credibility could be affected by the review design. In the movie industry, the most common review designs are reviews consisting of text with or without a star rating. Those movie reviews often are written by users (consumers) or experts (people with specifically movie knowledge). A more invisible part of the review however is equally important. The review content could be perceived as corresponding or contradictory, as a result of the use of the review with star rating design. In addition, research shows that depending on the valence of WOM to what extent the WOM is perceived credible, the impact on the decision is influenced (Fiske, 1980; Arndt, 1967). With this, the valence of the online movie review could influence the perceived review credibility, this will be explained further.

2.4 Review valence

The reason why the content of online reviews is mostly to classify on valence (having positive or negative direction) (Chatterjee, 2001), is because the motivation of consumers for writing a review is often a result of their previously too high or too low product expectations (Bone, 1995). According to Cheung and Thadai (2012) review valence is one of the most important factors with respect to the response. In addition, earlier studies have proved that the valence of information could have a different influence in WOM communications (Kanouse, 1984; Weinberger, M.G., Allen, C.T., and Dillon, W.R., 1981). Furthermore, the main effect of relatively negative online reviews is that they decrease the opportunity of purchasing because

consumers are less inclined to purchase products that have negative associations (Basuroy et al., 2003). Because of this, the frequency of new online reviews of this product will decline and the average opinion will not be changed.

Because of the above-mentioned current findings, it is key to take influential online review components into account when aiming at obtaining more realistic results. The influence of negatively framed information is stronger compared with positively framed information during the WOM communications. This could be explained by many reasons. One of the reasons is that negative messages could lead to a reduction of the possibility that product information is actually sent by someone. Another reason could be the natural tendency of people to avoid taking risks. The negative WOM could be interpreted as helping tool to avoid people from making a wrong purchase decision (Chevalier & Mayzlin, 2006).

Research results suggest consumers' opinions with respect to products are malleable by product reviews (Sorensen & Rasmussen, 2004). Results of Reinstein and Snyder (2000) study about movie reviews find relatively large influence effects of positive reviews on sales. But they could not find evidence for the increase of the sales after reading negative reviews. East, Hammond, and Lomax (2008) determined the different effects of PWOM (positive word-of-mouth) and NWOM (negative word-of-mouth). In general, PWOM has more impact on the decisions than NWOM. In addition, many studies suggest that a positive review has a larger impact on consumer behaviour than a negative review (van Dam, 2006; Vermeulen & Seegers, 2009). Furthermore, research of Basuroy et al (2003) suggests that the box office revenues significantly decrease by the use of negative reviews. There are many different studies that deliver evidence which indicated that purchase intentions of potential buyers are influenced more by NWOM compared to PNOM information (Brown and Reingen 1987; Weinberger, Allen and Dillon 1981). The attention will be greater for negative information compared to positive information. Because of this, the NWOM will have a stronger effect on the purchase intention in comparison to PWOM (Fiske, 1980; Arndt, 1967).

According to Huang and Chen (2006), the trustworthiness of the original advertising can be reduced by the received negative information. Research about unfavourable product ratings suggests that in case information about a firm or object received from recommendations or opinions of others, negative information may have perceived as more generalizable and credible than positive information. The suggested reason for this is because there is no underlying interest to share negative information, WOM sources have nothing to gain (Mizerski, 1982). In

addition, existing study shows that consumers attach more value to negative information compared to positive information (Skowronski and Carlson, 1987). Which means that in case an online review consists of negative information is perceived as more valuable. Based on above-explained research findings, the following hypothesis is expected:

Hypothesis 1.1: $W2W_P > W2W_B$

Hypothesis 1.2: $W2W_B > W2W_N$

Hypothesis 2: $W2W_P - W2W_B < W2W_N - W2W_B$

Hypothesis 3: $CR_P < CR_N$

2.5 Star rating and written reviews

During the process that consumers obtaining information about the certain experience of products, it is proved that the ins and outs of eWOM have to be understood to get the intended effects (Sparks & Browning, 2011). Other influential effects of helpfulness of online reviews are the design and the content of the message of online reviews (Zhou & Guo, 2017; Agnihotri & Bhattacharya, 2016; Schindler and Bickart, 2012). These studies proved that information refers to the reviewer and the non-evaluative product information is associated with the helpfulness of the reviews. The elements that make the reviews more entertaining are associated with more valuable reviews. The stylistic elements that damage the level of clarity of the reviews are associated with less valuable reviews (Schindler and Bickart, 2012). This finding underlines the importance of the design of the online reviews. The movie industry used online reviews that consist of two different parts: rating and written reviews. Because of the importance of review layout, it will be necessary to take these different review parts into account.

Based on representative sample of 1,031 US-based consumers, 54% of consumers pay attention to and consider the average star rating of online reviews, which make star rating to be the most important factor (BrightLocal, 2017). Although the mentioned research is not scientifically well-founded, it suggests that star rating could play a key role in the review interpretation. Already mentioned research about the influences of WOM on the sales of online books proved that online rating reviews have an influence on individuals' decisions making processes

(Chevalier & Mayzlin, 2006). This research was based on the online ratings of two different websites: bn.com and amazon.com. It indicates that the difference in sales for a specific book was related to two different main factors: the average star-rating of the books and the number of reviews on each website. The research also indicated that books with a relatively higher score of star-rating (independently of the written text) will have relatively higher sales. The reason for this sales increase could be based on the higher predicted product desirability as result of the higher star-rating. This could finally have influence on the strengths of the first mentioned effect on the prediction of product desirability. Alternatively, when people see the given star-rating before they read the written text of the review, they could be interpreting this written text in another way. In other words, the given star-rating could push the consumer to emphasise contents of the written text with a specific direction, either positively or negatively. This is based on the paradigm of the 'anchoring effect'.

This effect occurs when people make an estimation based on the first available information or reference, also called 'anchor'. The reference or anchor could be designed as a number, amount, prices, symbols and other alternatives which are easy to interpret. Using anchoring make possible to create a cognitive bias that indicates the common tendency of people to rely too heavily on the delivered reference. In case the reference is not relevant or in contrast with other information, it manipulated final judgments of human, owing to the adjustments people do to interpreting other information around the reference (Tversky and Kahneman, 1974). Furthermore, this heuristic was examined by Jacowitz and Kahneman (1995). By explaining the study of Jacowitz and Kahneman, it will give an example of what anchoring effect is. First, Participants were asked to make a comparative assessment by answering the question: 'Is the population of Chicago more or less than 200,000?'. Secondly, participants had to make an absolute estimation when they were asked: 'What is the actual population of Chicago?'. The received answers proved that people tend to think Chicago is densely populated after considering whether its population is more or less than 5 million in comparison with people who considering whether its population is more or less than 200,000 (Epley and Gilovich, 2006; Jacowitz and Kahneman, 1995).

Although one might think that consumers only use the star rating as evaluation, two-thirds of the online population want to take time to read the review text as well (69% for positive review and 63% for negative review), which is based on non-scientific research with respect to 1046 respondents who have had experiences with the customer service of a mid-sized company

(Dimensionalresearch, 2013). Online movie reviews mostly consist of a star rating and text. It should be clear in this case that the star rating design could play the anchor role during the considerations people make, by interpreting the total online movie review. It would therefore be interesting to

measure these anchoring effects in the movie industry and to determine whether this influences the willingness to watch a movie and as well as the perceived review credibility. Based on the research findings from the Chicago experiment (Epley and Gilovich, 2006; Jacowitz and Kahneman, 1995), the extent of willingness to watch a movie will be more or less the same in case the star rating and the text fit. The certain star rating valence (positive ≥ 4 stars and negative ≤ 2 stars) will not have a reinforcing effect on the willingness to watch a movie in case the star rating and text fit with respect to the willingness to watch a movie by using the same review text (e.g. the willingness to watch movie will not have a higher score if the review consists of a positive star rating as well as text). On the other hand, according to anchoring and adjusting effect, the willingness to watch a movie will have a score related to the used star rating, in case a review consists of a star rating and text with no fit. With this, it is expected that the use of incongruent reviews (reviews without matching star rating and text valence) will have a significant differential effect on the willingness to watch a movie compared to no review use. In addition, the anchoring effect will provide similar willingness to watch outcomes to the reviews with internal matching reviews (the comparison between incongruent and congruent reviews but both the same star rating). Latter mentioned provides the basis for the following formulated hypotheses:

Hypothesis 4: $W2W_{PP} = W2W_P$; $W2W_{NN} = W2W_N$

Hypothesis 5: $W2W_{PN} > W2W_{_N}$; $W2W_{NP} < W2W_{_P}$

$W2W_{PN} > W2W_{NN}$; $W2W_{NP} < W2W_{PP}$

Furthermore, Research of Basuroy et al. (2003) proved that negative effects of a relatively poor review are significantly greater than positive effects of a relatively good review. It should be taken into account that this finding only covers 'the opening week' (refers to the first week of a film's run), which means that there was probably no time to react or counter negative reviews as well as positive reviews. In addition, the context of online news messages has an impact on

the evaluations of news credibility (Thorson, Vraga, and Ekdale, 2010). This study suggests that the content of a message will be assigned a higher credibility in case the context is the opposite, for instance, uncivil context. Furthermore, if the anchoring paradigm affects the willingness to watch a movie, it could affect the perceived credibility as well. For instance, people interpreting the star rating first, could allow this rating to be the frame of reference in the human's mind. Even if the star rating does not relate to the review credibility, people could be manipulated by seeing the anchor (number of stars) (Tversky and Kahneman, 1974). However, inconsequent or contradicting substance could generate suspicion with the reader as well. The contradictory messages of the online movie review will be interpreted as more complex in comparison to an online movie review with a consistent content. People have a tendency to regard the consistent and simple to interpret content as more credible relative to the more complex content (Pearl, 1978). For instance, when people first see a relatively high star-rating and read a relatively negative written review part, the perceived credibility of the total online review could be reduced.

Based on the above-mentioned prior research outcomes and the descriptions of the possible important relations between these different outcomes, the following hypotheses are formulated:

Hypothesis 6: $CR_{PP}, CR_{NN} > CR_{PN}, CR_{NP}$

2.6 Review source

Nowadays, consumers are able to collect not just expert reviews but user reviews as well, which may lead to the consumer taking several more factors into account than expert review content alone. Recent studies suggest that the expectations related to the communicators, intrinsic motives and preoccupations, have a key role during valuations of online reviews made by consumers and probably affect review impact (Sen & Lerman, 2007; Vermeulen, Das, & Swager, 2008).

A previous study has focused on the relation of consumer's review and the marketing communication mix (Chen and Xie, 2008). The authors found when and how the two different type of information sources, consumers as well as sellers, could be used as a marketing communication element. Based on these findings, marketing strategy could be improved by using reviews and recommendations to increase the level of efficiency and effectiveness during

the customer decision journey. This refers to the distinctions between self and others, in-group and out-group members, similar and dissimilar others who are all instances of social distance (Trope, Liberman, and Wakslak, 2007). Furthermore, the effects of social and temporal distance on consumer's recommendations to peer recommendations are different. Research determined that 'distant others' do not always have a smaller impact than 'close others'. Recommendations from 'distant others' are more influential than those from 'close others' in shifting distant-future preferences. To shifting preferences in the near-future, the 'close others' are more influential than 'distant others' (Zhao and Xie, 2011).

If we implement this to the user and expert groups, users belong the most to 'close others', because consumers see themselves as 'potential users' instead of experts. Meaning that expert reviews are mostly more influential than user reviews in shifting distant-future preferences. In addition, user reviews are mostly more influential than expert reviews in shifting near-future preferences. Latter mentioned is mostly related to the online movie reviews because in all likelihood people search for online movie reviews when they want to watch a movie.

The main advantage of online user reviews is, people have a high level of experiential knowledge and information that can easily be made available to many users. In spite of some people lacking official expertise in certain domains, others may consider their experience or knowledge on the topic as having lots of experiential credibility's (Flanagin & Metzger, 2008). The core advantage of using expert reviews is, it creates certain credibility as result of the specific product or service view, which arises from experiences with and knowledge about several similar products or services.

Eliashberg and Shugan (1997) found the two roles that an expert can have. These roles are the role of influencer and the role of predictor during the decision making process. Klucharev, Schmidts and Fernandez (2008) observed a significant correlation between the perceived degree of expertise of a celebrity regarding a product and the anterior cingulate cortex (this part of the brain is responsible for cognitive functions, e.g. decision-making). The significant correlation implies that the greater the perceived expertise, the greater the intention to purchase the product. This effect was the same remembering the product. Altogether, they concluded that expert reviews lead to a 12% increase more favourable attitude toward the product. This may eventually lead to a decreased risk of purchasing an unknown product. With this, an unknown movie will have more willingness to watch in case expert reviews were the foundation for the movie decision, compared to user reviews.

Research shows that seller-created information is likely to be less credible than consumer-created information. The reason for this is that the credibility of information is often positively related to the trustworthiness of the source (Wilson and Sherrell, 1993). The credibility of expert reviews hinges not on the aggregated opinions of many but on traditional markers of expertise (Andrew and Metsger, 2013). In contrast to the user reviews, expert reviews increase the level of perceived credibility that is based on the suspected extra interests. These extra interests could be derived from reputation and responsibility interests. In addition, the use of the elaboration likelihood model led the authors to conclude that the credibility of a source acts as a peripheral cue. Consumers expect from experts to be usually correct, just because of their expert title (Petty and Cacioppo, 1986).

However, a previous study (Willemsem, Neijens and Bronner, 2012) reveals that experts having less trustworthiness than laypersons (users), despite the experts being perceived as having more expert knowledge. The second part of the same study made clear that first mentioned only occurs when the expert status is based on self-claims. In case the expert status of the reviews was based on peer ratings, the review source was assessed as having both trustworthiness and expertise (Willemsem, Neijens and Bronner, 2012). Other studies indicated that consumers assign more value to consumer's recommendations in comparison to recommendations written by professionals. As appears from these studies, a consumer perceives opinions from other consumers' to be more neutral and less biased. In addition, according to these studies consumers relate more easily to consumers' experiences (Bickart & Schindler, 2001). In contrast, more previous research suggests that reviews are written by experts, influence consumers more. The reason for consumers perceiving experts as more reliable and better informed is related to the increasing experience with online reviews (Senecal & Nantel, 2004).

Based on the prior research findings, the following phenomenon is expected:

Hypothesis 7: $W2W_{PNE} < W2W_{NPE}$; $W2W_{PNU} > W2W_{NPU}$

Hypothesis 8: $CR_E \neq CR_U$

Based on the above-mentioned expectations that star rating has more impact in case the review is written by a user and review text has more impact in case the review is written by an expert, the following expectation is derived.

$$\text{Hypothesis 9: } CR_{PPE} - CR_{PE} < CR_{PPU} - CR_{PU}; \\ CR_{NNE} - CR_{NE} < CR_{NNU} - CR_{NU}$$

$$\text{Hypothesis 10: } W2W_{PPE} < W2W_{PPU}; W2W_{PE} < W2W_{PU} \\ W2W_{NNE} < W2W_{NNU}; W2W_{NE} < W2W_{NU}$$

Given the different expected effects of the different review components, it is necessary to measure the coherent effects. The most necessary reason for this is that this also plays a part in reality. There are no existing studies that test these components at the same time. To elaborate on the already mentioned research findings, the expected effects will be explained as follows. When we assume the review reader will take underlying source interests into account, it is expected that there is impact difference between review star rating and text relative to the two different sources. The quality of the review text will have more impact on the willingness to watch a movie in case the review is written by an expert compared to a user review. This is because experts could be paid for giving their assessments and at the same time, they will be judged on these assessments. For user reviews, this does not apply. Furthermore, when the online review is positive and written by a user, it will have a higher perceived credibility compared to an online review written by an expert.

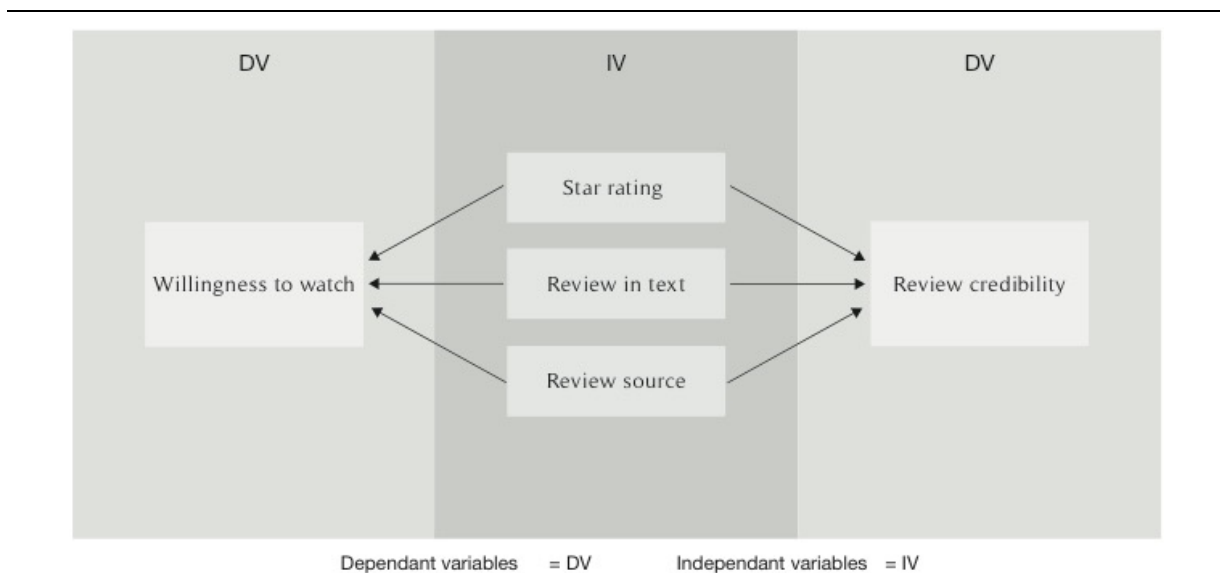
This last mentioned is because users do not have to deal with extra interests in general in contrast to an expert who may have career-related interests. The review reader, may take this into account but also in opposite case (negative online reviews written by experts will have higher perceived credibility compared to negative online reviews written by users). The above-mentioned assumptions underlie the following hypotheses:

$$\text{Hypothesis 11: } W2W_{PNE} < W2W_{PNU}; W2W_{NPE} > W2W_{NPU} \\ W2W_{NPE} = W2W_{PNU}; W2W_{PNE} = W2W_{NPU}$$

2.7 Conceptual Framework

Briefly, this study has the aim to examine the following effects: (a) Effect of the use of online movie reviews relative to the willingness to watch a movie; (b) Separate effects of the review valence, review internal match and review sources on relative to the willingness to watch a movie; (c) Separate effects of the review valence, review internal match and review sources on relative to the perceived review credibility; (d) The combination effect of review design (star rating and text) and source relative to the willingness to watch a movie; (e) The combination effect of review valence and source relative to the perceived review credibility.

Figure 1.
CONCEPTUAL FRAMEWORK



3. Research Methodology

After the first two sections, the introduction and the theoretical framework, the next section consists of a detailed description of the research method and design, which starts with 3.1 Research design, followed by 3.2 Sample, 3.3 Experiment overview, 3.4 Measures, 3.5 Manipulation check and 3.6 Statistical methodology.

3.1 Research design

This research used the movie industry as an empirical setting. In reality, consumers have to rely on the existing evaluations of other consumers. Due to this, the used movie reviews are existing movie reviews which are collected from an existing movie review platform ‘metacritics’. In order to increase credibility of the impression people have of movie reviews, a pre-survey is conducted. This pre-survey showed the participants some existing movie reviews and content related questions (Appendix 1 and 2). The outcomes are the basis of the selection of the movie reviews for the final survey (Appendix 3). Furthermore, the movie reviews consisted of (expert) reviews and user reviews, both reviews types were expressed by a review text and/or a star rating. But to measure the effects of different ratings on the dependent variables, it was necessary to manipulate the given star-rating combined with the same written review text. Based on the mentioned findings, the star rating, as well as the review text of the conducted survey, expressed a positive valence of four-stars or a negative valence of two-stars. The pre-survey provided the right selected review text, which was interpreted as four-star or two-star ratings. All movie related questions started with general movie information. This consisted of a fictional movie title, short movie content summary, movie genre(s) and runtime. All survey conditions showed several of the same general related questions in the end (Appendix 3).

3.2 Sample

To test the formulated hypotheses an experimental approach was used. This approach was based on an online questionnaire that was sent out to the participants via AmazonTurkey. Participants were presented in the United States during answer the survey online. The motivation for participating in this survey was the possibility to earn money, which was based on the current minimum wage. During this study, the participants were randomly assigned to one of the survey conditions. Table 2 shows the average profile of the participants of this study.

Table 2.
PARTICIPANTS PROFILE

Gender	Male 292 Female 228
Average Age	Between 25 - 44
Average Review Usage	Half of the time
Average Review Opinion	Between helpful and somewhat helpful
Most Preferred Movie Genre	Comedy Movies
Least Preferred Movie Genre	Art House Movies

3.3 Experiment overview

The survey started with an introduction which consisted of information about the research goal, practical participation information and guarantee of careful handling of the given answers. After reading the introduction, participants were randomly assigned to one of the 13 survey conditions. The survey conditions that were used are displayed in table 3. The participants did not know there were more survey conditions. Regardless of which survey condition, participants had to read some general information about a movie. Dependent on the survey condition, this general information was supplemented with a movie review. This movie review could consist of a star rating and text or just text. All used movie reviews were contained a source. After reading the general information or movie review, the participants were asked to what extent they wanted to watch the movie, followed by the question about how credible they think the movie review was. The last mentioned question was only present in the survey conditions with movie reviews. All survey conditions contained two different movies, with related abovementioned questions. After the movie related questions, participants had to answer general questions. The first question was about his or her opinion about movie reviews in general. The second question was about his or her movie review usage in the current situation. This question was followed by a question where the participant had to rank six different movie genres based on his or her preferences. The last two questions of the whole survey were about

the participants themselves, such as age and gender. All questions had to be answered in order to complete the survey.

Table 3.
SURVEY CONDITIONS

Conditions	Sub Conditions	General Movie Information	Movie Review	Star Rating	Text Valence	Source
Baseline						
	Baseline	Yes	No	–	–	–
Congruent	PPE	Yes	Yes	4	Positive	Expert
	PPU	Yes	Yes	4	Positive	User
	NNE	Yes	Yes	2	Negative	Expert
	NNU	Yes	Yes	2	Negative	User
Incongruent						
	PNE	Yes	Yes	4	Negative	Expert
	PNU	Yes	Yes	4	Negative	User
	NPE	Yes	Yes	2	Positive	Expert
	NPU	Yes	Yes	2	Positive	User
Only Text						
	_PE	Yes	Yes	No	Positive	Expert
	_PU	Yes	Yes	No	Positive	User
	_NE	Yes	Yes	No	Negative	Expert
	_NU	Yes	Yes	No	Negative	User

3.4 Measurements

The first two questions after reading the movie review or after movie general information, are measured on a Likert scale. The two questions had different Likert scale numbers of Likert scale. The existing several types of research conflict on the optimal numbers of Likert scale (Adelson & McCoahc, 2010). Therefore, this study used two different numbers of Likert scale (11-5). The numbers of Likert scale are based on the extent to which the nuance of answer options was meaningful, in which was related to the content.

Willingness to Watch. After every movie review or in the case of the baseline condition after general movie information, the participants were asked to what extent they want to watch the movie. To answer this question, they had to express their willingness into an amount of points. An 11-scaled answer option was given, in which 0 point was ‘Definitely Not’ and 10 points indicated ‘Definitely Yes’. 5 points were related to a neutral willingness.

Review Credibility. After reading the movie review and answering the question about to what extent the reader wanted to watch the movie, participants had to give their opinion about the movie review itself. In order to find out how credible the certain movie review was considered to be the following question was asked: ‘How credible was the online movie review?’. The participants could answer this question by choosing one of the following answer options: a) Not at all credible, b) Not credible, c) Somewhat credible, d) Credible and e) Very credible.

Review Valence. To measure the different review valence effect on the two dependent variables, there were two survey conditions which provided this probability. Conditions: Congruent and Only Text, consisted of completely positive or negative movie reviews. Because this, the different review valence effect on ‘the willingness to watch a movie’ as well as ‘the perceived review credibility’ could be measured.

Matching – No Matching. In contrast to previously mentioned conditions, the ‘Incongruent’ survey condition was needed. to measure the different effects of matching and non-matching reviews on the two different dependent variables. The ‘incongruent’ survey condition consisted of non-matching movie reviews which means that in case the star rating was relatively positive, the review text was negative and vice versa. When you compare the outcomes of the ‘incongruent’ survey condition with the outcomes of the ‘congruent’ survey condition, you measure the different effects of matching and non-matching reviews. Worthy to note here is that the outcomes of survey condition ‘Only text’ was not used, because it could be a less

trustworthy comparison due to the absence of star rating. The comparison between the outcomes within survey condition 'Incongruent', of the sub-conditions 'PNE, PNU, NPE and NPU', provided the opportunity to find the different dominance of the star rating and review text.

Review Source. All used movie reviews contained a source reference, during the conducted survey. This allowed measuring the effects of the different sources use on the two dependent variables. The different effects of the two sources 'Expert' and 'User' could be measured in three survey conditions 'Congruent', 'Incongruent' and 'Only Text'.

3.5 Manipulation check

According to Mudambi and Schuff (2010), the review could have a more convincing effect when consumers used longer online reviews. To eliminate the possible manipulations effects of text quantity, the used movie reviews had the same amount of words. In order to obtain a confirmation whether there is no manipulation during the questionnaire, a pre-test of the final survey was performed. In addition, this pre-test had to clarify that the participants would not spend too much time on answering the questions and the reviews that are used will not be misleading or too obvious. Furthermore, the current study took the following ethical issues into account: a) Anonymity, b) no damage to participants, c) participation is fully voluntary and d) confidentiality (Baxter & Barbie, 2004). The main reason for this is to receive as reliable outcomes as possible. The first survey page, contained information to clarify to participants that they had to answer the questions voluntarily, that they had the right to stop anytime and their collected data will be handled anonymously.

3.6 Statistical methodology

The collected data of the survey is analysed using SPSS software after being collected through Qualtrics. The implementation of several one-way ANOVA tests made it possible to measure the different effects of the independent variable on the dependent variables. All ANOVA results were based on a confidence level of 90% and a significance level of 0.1. This is because the sample sizes of the 13 survey conditions were relatively small (33 – 45 participants), which could lead to a relatively high effect size (Ferguson, 2009; Bakeman, 2005; Cohen, 1992). During the execution of the survey, there were 13 different groups: a control group and 12

experimental groups. The respondents that belong to the control group received a survey with questions of the baseline condition. The respondents received a survey containing questions based on adjusted reviews. With this, the used statistics are based on a between-subject design.

Chapter 4 | Results

In the following chapter, the research data will be described by discussing the empirical results of the path analysis. The statistical software tool SPSS 23 was used for testing the survey data. Because of the present large effect size, tests with a $p < .1$ were determined as statistically significant and tests with $p > .1$ were determined as not statistically significant. The formulated hypotheses will be treated content chronologically instead of consecutive hypothesis number.

4.1 Assumptions ANOVA

- All cases regarding the 13 conditions were independent and selected randomly
- All 13 conditions were tested by at least 30 cases, which fit the main rule of at least 20 respondents and make testing the normal distribution of the variables unnecessary (List et al., 2011; Khan & Rayner, 2003).
- The dependent variables (Willingness to watch and credibility of review) are measured at the continuous level.
- By applying the Levene's test, it has turned out for all conducted ANOVA tests that there was the homogeneity of variances (with $p = .665$ and $p = .692$).

4.2 Analysis of variance

The analysis of variance or ANOVA is a test for the relation (with unilateral influence) between an independent categorical variable and a dependent interval-scaled variable. With the analysis of variance, more than two groups of the categorical variable (the factor) will be distinguished. In addition, the population mean of the independent groups will be tested whether those are equal. With this, the analysis of variance will answer the question whether a categorical variable statistical significant affects an interval-scaled variable (Ghosh & Vogt, 2012; Janssens, Wijnen, De Pelsmacker & Van Kenhove, 2008).

To confirm or to reject the hypotheses of this research, the following independent variables were tested: Star rating, review text and review source. The star rating was tested using three conditions: high rating (4 stars), low rating (2 stars) and no rating. All in the surveys used reviews consisting of a relatively positive text or relatively negative text and in addition, all

reviews showed one of the two following sources: Expert and User. Therefore, the applied ANOVA tests answer the question whether these three mentioned independent variables affect the dependent variables, Willingness to watch a movie (willingness) and review credibility (credibility).

4.3 Main effects ‘Willingness to watch’

4.3.1 Review use

The willingness to watch a movie is expected to be higher by using positive reviews (positive star rating and text or only positive text) compared to no use of reviews. The applied ANOVA (Appendix 4, Table 5.1) indicates that there is no significant difference in willingness to watch a movie by using positive reviews (PPE, PPU, _PE and _PU) compared to no use of reviews (Baseline). As a result, hypothesis 1.1 has to be reject, which is as follows: ‘The usage of positive online movie reviews has a significantly higher willingness to watch a movie compared to non-use of online movie reviews’.

In addition, it is expected that the willingness to watch a movie will be different by using negative reviews (negative star rating or only negative text) compared to no use of reviews. The main value of different willingness to watch a movie between negative reviews (NNE, NNU, _NE and _NU) and baseline is not significantly different (Appendix 4, Table 5.2). However, this could be different if you examine the sub conditions separately. This time ANOVA (Appendix 4, Table 5.5), indicated that there is a significant difference in willingness to watch a movie by using negative reviews compared to no use. The ANOVA outcomes show that the willingness to watch a movie will be lower in case the used online movie review is negative compared to non-use of online movie review (BASE and NNU: $MD = 1,6509$, $P = ,001$, BASE and NNE: $MD = 1,0423$, $P = ,039$ and BASE and _NE: $MD = 1,3215$, $P = ,008$). Consequently, that support is found for the last part of hypothesis 1.2.

Although, the ANOVA results about the main effect are not significantly different (Appendix 4, Table 5.3), it is expected that using negative reviews with only negative text or and with a negative star rating, have a higher impact on the willingness to watch a movie compared to the use of positive reviews with only positive text or and positive star rating. With the above-mentioned findings (hypothesis 1.2), hypothesis 2 is about the difference in effect size between positive and negative reviews, where it is expected that negative reviews have a higher impact

on the willingness to watch a movie than the use of positive reviews, is confirmed. Noteworthy is that using reviews with only negative text and written by users have no significant difference in willingness to watch a movie compared to no reviews use (BASE and _NU: $p = .104$).

4.3.2 Rating and written review

It is expected, that adding a star rating to review text will have no significant difference in willingness to watch a movie when the star rating fits the valence of the review text compared to no use of a star rating. As expected, the results of the applied ANOVA (Appendix 4, Table 5.5) indicate that there is no statistically significant difference in willingness to watch a movie by adding matching star rating (PPE and _PE: $P = .906$, PPU and _PU: $P = .164$, NNE and _NE: $P = .596$ and NNU and _NU: $P = .11$). This means support was found for hypothesis 4.

Another type of influence that adding star rating could have on the willingness to watch a movie is the anchoring effect. It is expected that adding a contradictory star rating to a review text (e.g. positive star rating to a negative text) will dominate the total review effect. According to the results of the ANOVA, in general, adding contradictory star rating has no significant different effect on the willing to watch a movie, which is unexpected (PNE and _NE: $P = .39$, NPE and _PE: $P = .251$ NPU and _PU: $P = .124$). Except in the case, where reviews with a negative text written by a user adding a positive star rating, has a significantly different effect on the willingness to watch a movie (PNU and _NU: $MD = 0.95$, $P = .063$). Hypothesis 5 part 1 is confirmed only by the review condition negative text written by user. Furthermore, using incongruent reviews (no matching directions between the star rating and text direction) is expected to have a significant difference in willingness to watch a movie compared to the use of congruent reviews with the same text valence and opposite star ratings. According to the outcomes of the ANOVA, part 2 of hypothesis 5 is supported by two cases, namely positive star rating and written by a user and negative star rating and written by a user. This means that there is no significant difference in willingness to watch a movie between the use of incongruent and congruent online movie reviews written by an expert. In case the online movie review is written by a user, there is a significant difference in willingness to watch a movie when the star rating has the opposite valence (NPU and PPU: $MD = -1.5227$, $P = .004$, PNU and NNU: $MD = 1.8030$, $P = .001$). This means that the expected effect of the incongruent valence of the star rating on willingness to watch a movie is supported in case the reviews are written by users.

The anchoring effect plays a key role in case the reviews are written by users. This confirmed the expectations of the role of star rating which is dominant in user reviews and the role of review text which is dominant in the expert reviews.

Based on the current findings, which confirm the anchoring effect only in cases where the review is written by a user, it would be relevant to measure the levelling effect. In case the incongruent online movie reviews have no significant difference in willingness to watch a movie compared to no review use, it could be mean that the participants took the valence of the star rating as well as the review text into account. The consequence of this phenomenon will be a levelling effect which is a middle score. This middle score will not be a significant difference with the baseline score of the willingness to watch a movie.

The ANOVA findings confirmed this expectation by three conditions (BASE and PNU: $P = 0,758$, BASE and NPE: $P = 0,347$ and BASE and NPU: $P = ,114$). Remarkably, reviews with a positive star rating, a negative text and written by an expert have a significant difference in willingness to watch a movie compared to no review use (BASE and PNE: $MD = 0,8923$, $P = ,061$). Noteworthy, the founded significant difference in willingness to watch a movie confirmed the absence of anchoring effect because the negative text of the review had a higher impact compared to the positive star rating because the willingness to watch a movie was lower compared to the baseline score. Based on the above mentioned findings, incongruent online movie reviews have a levelling effects, which means that the participants take the star rating as well as the review text into account and compensate the opposite valences in one review.

4.3.4 Source

According to the ANOVA results with respect to the main value main effect of an expert review and a user review on the willingness to watch a movie is not significantly different (Appendix 4, Table 5.4). However, according to the results of the ANOVA test based on the sub conditions (Appendix 4, Table 5.5), the use of reviews has as expected a bigger chance to make significant difference in the willingness to watch a movie when it is written by an expert, because three out of four significances results are written by an expert (BASE and PNE: $P = ,061$, BASE and NNE: $P = ,039$ and BASE and _NE: $P = ,008$). On the other hand, using reviews with a negative star rating, negative text and written by users have a higher impact on the willingness to watch a movie (BASE and NNU: $MD = 1,6509$, $P = ,001$) compared to negative reviews written by experts (BASE and NNE: $MD = 1,0423$, $P = ,039$ and BASE and _NE: $MD = 1,3215$, $P = ,008$).

Based on the current findings of anchoring effects, it is confirmed that a star rating has a dominant role with respect to review text in the evaluation of the total user review. But the dominant role of the review text in the case the review is written by experts is not measured. Based on the ANOVA findings (Appendix 4, Table 5.5) there is no statistically significant difference in effect on the willingness to watch a movie between star rating and text in case expert reviews is used (PNE and NPE: $P = ,369$). When the review is written by users, there is a statistical difference in effect on the willingness to watch a movie between star rating and text, whereby review star rating has a dominant role (PNU and NPU, $MD = 0,9167$, $P = ,069$). Therefore, hypothesis 7.1 ‘During interpret an expert review, the review text has a dominant role in willingness to watch a movie in comparison of a star rating’ is not supported by the statistical findings. However, for hypothesis 7.2 ‘During interpret of a user review, the star rating has a dominant role in willingness to watch a movie in comparison of the review text’ support was found. Moreover, when we compare the different sizes of impact on the willingness to watch a movie, of negative and positive review texts written by expert and written by users, we can conclude that the review text written by experts has more impact because there is a significant difference in willingness to watch a movie between positive review text and negative review text both written by experts. At the same time, there is no significance in willingness to watch a movie between positive and negative review texts written by users. ($_PE$ and $_NE$: $MD = 1,4282$ $P = ,005$ compare to $_PU$ and $_NU$: $P = ,113$). This means that support is found for hypothesis 7.1 in case the review consists of only text.

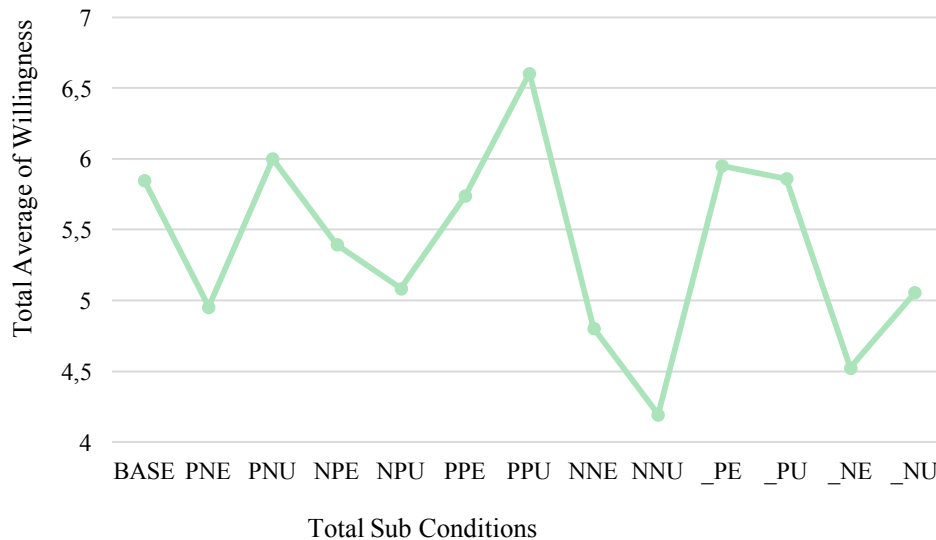
In addition, hypothesis 10 formulates the expectation that positive reviews written by a user will have a significantly higher willingness to watch a movie in comparison with positive expert reviews. In case the review is negative it is expecting that the willingness to watch a movie is lower in case it is written by an expert compared to a user review. Worthy to note here is that there is no significant difference was found by the ANOVA (Appendix 4, Table 5.4) in willingness to watch a movie between the two different sources and the positive and negative reviews. This is an unexpected finding, as negative reviews written by an expert could have more impact compared to negative user reviews based on their expert interests. Experts may have to take more risks by writing a negative review because this could damage their position in the market. At the same time, users don’t have an extra interest which could play a role when writing positive reviews.

According to the results of ANOVA with respect to the sub conditions (Appendix 4, Table 5.5) hypothesis 10.1 and 10.2 are not supported by the statistical findings (PPE and PPU: $P = ,257$, _PE and _PU: $P = ,484$, NNE and NNU: $P = ,265$, _NE and _NU: $P = ,308$).

Furthermore, hypothesis 11 is about the different impacts of the review star rating and review text across the different sources. Hypothesis 11 formulated the expected significant different impact on willingness to watch a movie in case of reviews with positive star rating and negative review text written by a user and an expert compared to reviews consist of a negative star rating and positive review text written by a user and an expert. In addition, because of the different expected dominant roles of star rating and review text with respect to the two sources: expert and user, the following effects are expected: 1) A movie review contained of a positive star rating and a negative review text written by a user and a movie review contained of a negative star rating and a positive review text written by an expert will have a higher willingness to watch a movie compared to 2) a movie review contained a negative star rating and a positive text written by a user and a movie review contained of a positive star rating and a negative review text written by an expert. The conducted ANOVA shows a significant result occurs in case of a review containing a positive star rating and negative text and no significant result occurs in case of a review containing a negative star rating and positive review text (PNE and PNU: MD = -1,0444, $P = ,036$, NPE and NPU: $P = ,531$, NPE and PNU: $P = ,229$, PNE and NPU $P = ,793$). The statistical results exclude the rejection of hypotheses 11.

Figure 2.

MEANS PLOTS: WILLINGNESS AND TOTAL SUB CONDITIONS



Baseline; BASE= No review;

Incongruent; PNE = Positive star rating, Negative text and Expert; PNU = Positive star rating, Negative text and User; NPE = Negative star rating, Positive text and Expert; NPU = Negative star rating, Positive text and User.

Congruent; PPE = Positive star rating, Positive text and Expert; PPU = Positive star rating, Positive text and User; NNE = Negative star rating, Negative text and Expert; NNU = Negative star rating, Negative text and User.

Only text; _PE = No star rating, Positive text and Expert; _PU = No star rating, Positive text and User; _NE = No star rating, Negative text and Expert; _NU = No star rating, Negative text and User.

4.4 Main effects 'Credibility'

As mentioned in the literature review, the expectations about the effects on the perceived review credibility are not directly related to the expected effects with regard to the willingness to watch a movie.

4.4.1 Valence

When people write a negative review it will have a higher perceived credibility compared to positive reviews, the reason for this expectation is explained in the literature review. As a result,

it is expected that there is a significant difference in credibility between positive and negative reviews. The main effect is not significantly different according to the ANOVA results (Appendix 4, Table 6.1). However, hypothesis 3 could be confirmed by analysing the difference in the perceived review credibility between the negative and positive reviews on the level of sub conditions. The results of the ANOVA with regard to the dependent variable: credibility (Appendix 4, Table 6.4) suggests that this hypothesis must be rejected and, more strongly, the opposite expectation effect has occurred in 3 out of 4 cases (PPE and NNE: $MD = 0,7297$, $P = 0$, PPU and NNU: $MD = 0,8485$, $P = 0$ and _PE and _NE: $MD = 0,2835$, $P = ,063$). There is no significant difference in credibility between a positive and a negative review without star rating written by a user (_PU and _NU: $P = ,527$).

Unexpectedly, the negative reviews have a lower perceived credibility compared to the positive reviews. As a result, hypothesis 3 ‘the negative online movie review has a significantly higher perceived review credibility compared to positive online movie review, is rejected.

4.4.2 Star rating

Adding a star rating is expected to increase the perceived credibility of the review in case the star rating is matching the valence of the text. This is because the star rating would underline the core message of the text and create a consistent message. Moreover, because of the expected dominant role of star rating in a user review and the dominant role of review text in an expert review, adding a matching star rating to a user review will have a significant higher perceived review credibility with respect to an expert review. The applied ANOVA (Appendix 4, Table 6.4) shows that this hypothesis is only confirmed when the review is positive and written by a user (PPU and _PU: $MD = 0,3322$, $P = ,041$). When the positive review is written by an expert it has no significantly different effect on the perceived review credibility (PPE and _PE: $P = ,308$). Remarkably, in the case the review is negative, it has the opposite effect. Meaning that when the negative review has no star rating it perceived a higher credibility compared to the negative review with a matching star rating. This effect is not different by using the different sources (NNE and _NE: $MD = -0,2961$, $P = ,064$ and NNU and NU: $MD = -0,4186$, $P = ,01$). Therefore, hypothesis 9 is rejected in case the review is negative and supported in case the review is positive.

4.4.3 Congruousness and Incongruousness

According to the current research findings which are mentioned in the literature review, the incongruent reviews (no match between star rating and text valence), will have a lower perceived review credibility compared to the congruent reviews (match between star rating and text valence). However, the statistical results of ANOVA with respect to the main effect do not confirm this expectation (Appendix 4, Table 6.2). On the other hand, the results of the ANOVA based on the sub conditions show that the hypothesis 6 is confirmed only in case with the incongruent reviews and positive congruent reviews (PNE and PPE: $MD = -0,3158$, $P = ,031$, PNU and PPU: $MD = -0,3834$, $P = ,019$, NPE and PPE: $MD = -0,4499$, $P = ,003$ and NPU and PPU: $MD = -0,4805$, $P = ,003$).

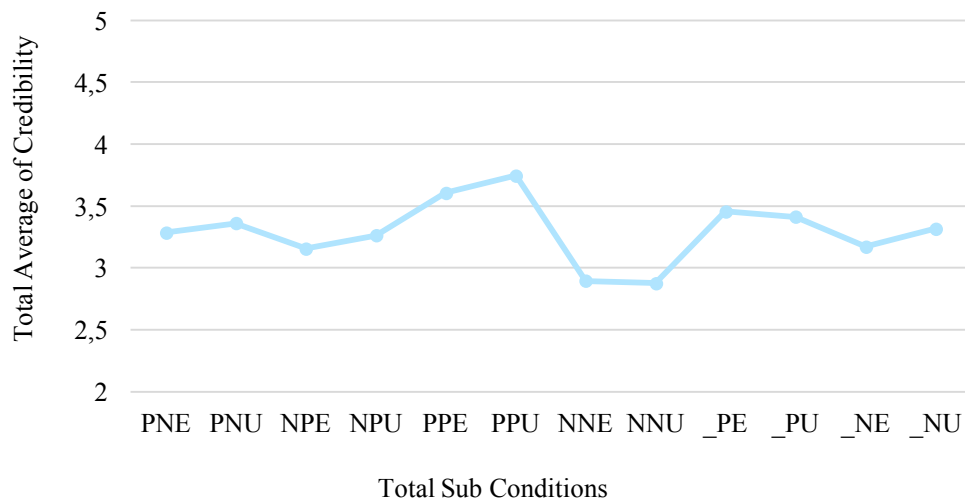
Worthy to note is that regarding incongruent reviews and congruent negative reviews the perceived credibility is significant different but unexpectedly in an opposite direction. The negative reviews have a lower perceived review credibility compared to the incongruent reviews (PNE and NNE: $MD = 0,4139$, $P = ,007$, PNU and NNU: $MD = 0,4650$, $P = ,004$, NPE and NNE: $MD = 0,2798$, $P = ,073$ and NPU and NNU: $MD = 0,3680$, $P = ,022$).

4.4.4 Source

Hypothesis 8 suggests that the usage of online movie reviews written by an expert has a significant different perceived review credibility compared to the usage of online movie review written by a user. The results of the ANOVA with respect to the main effect indicates that there is no significant difference in the perceived credibility between user and expert reviews (Appendix 4, Table 6.3). The outcomes of the applied ANOVA based on sub conditions showed in the second matrix (Appendix 4, Table 6.3), do not support this expectation (PPE and PPU: $P = ,386$, NNE and NNU: $P = ,909$, _PE and _PU: $P = ,769$, _NE and _NU: $P = ,363$, PNE and PNU: $P = ,641$, NPE and NPU: $P = ,474$). This means that hypothesis 8 must be rejected.

Figure 3.

MEANS PLOTS: CREDIBILITY AND TOTAL SUB CONDITIONS



Incongruent; PNE = Positive star rating, Negative text and Expert; PNU = Positive star rating, Negative text and User; NPE = Negative star rating, Positive text and Expert; NPU = Negative star rating, Positive text and User.

Congruent; PPE = Positive star rating, Positive text and Expert; PPU = Positive star rating, Positive text and User; NNE = Negative star rating, Negative text and Expert; NNU = Negative star rating, Negative text and User.

Only text; _PE = No star rating, Positive text and Expert; _PU = No star rating, Positive text and User; _NE = No star rating, Negative text and Expert; _NU = No star rating, Negative text and User.

4.3.4 Halo-effect

It has been found that the most of the hypotheses about the credibility are not confirmed by the ANOVA findings. In order to be sure that this study takes all underlying processes into account, deliver the right evidence and finally consist of a reliable conclusion, it is necessary to detect any halo-effects.

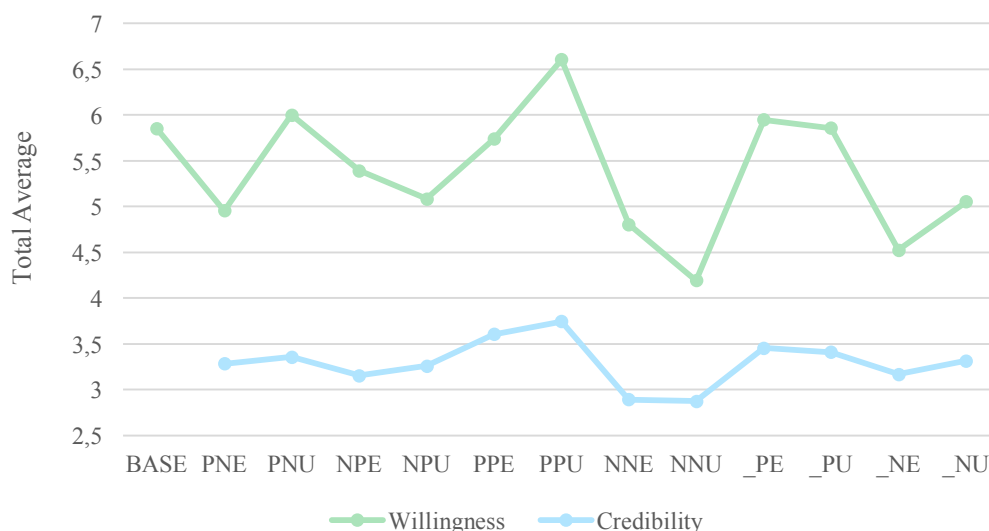
According to Allenby, Grilbride and Yang (2005), responses to questions in a survey can reflect a behaviour process that influences the multiple response items. In addition, the presence of a certain quality gives the respondent the suggestion that other qualities are present as well, what is called a halo effect. The consequences of the presence of halo effect in the survey responses are that it damages the reliability of the outcomes.

To detect the presence of a halo effect in the survey outcomes, it is necessary that there is an unexpected pattern of a relative equal response to different survey questions (Allenby and Rossi, 2003). This is what happened to the two different questions first about the willingness to watch a movie and second about the credibility of the review (Appendix 3, Final survey). Figure 4 shows that the outcomes of the two different questions have an equal pattern. It is necessary to take this finding into account with regard to the weight of the outcomes about the credibility.

Another explanation for the unexpected similar outcomes of the perceived review credibility with respect to the outcomes related to the willingness to watch a movie, is that participants are influenced by the review content valence. This suggests that participant's assessment of the review credibility affect the positive or negative direction of the review content.

Figure 4.

MEANS PLOTS: WILLINGNESS, CREDIBILITY AND TOTAL SUB CONDITIONS



Baseline; BASE= No review;

Incongruent; PNE = Positive star rating, Negative text and Expert; PNU = Positive star rating, Negative text and User; NPE = Negative star rating, Positive text and Expert; NPU = Negative star rating, Positive text and User.

Congruent; PPE = Positive star rating, Positive text and Expert; PPU = Positive star rating, Positive text and User; NNE = Negative star rating, Negative text and Expert; NNU = Negative star rating, Negative text and User.

Only text; _PE = No star rating, Positive text and Expert; _PU = No star rating, Positive text and User; _NE = No star rating, Negative text and Expert; _NU = No star rating, Negative text and User.

The mentioned outcomes of all hypotheses of the current study are mapped in Table 7.

Table 7.
HYPOTHESIS OUTCOMES

▪ Hypothesis 1.(1): $W2WP > W2WB$	▪ Rejected
▪ Hypothesis 1.(2): $W2WB > W2WN$	▪ Support was found (NNE, NNU and _NE vs. Baseline)
▪ Hypothesis 2: $W2WP - W2WB < W2WB - W2WN$	▪ Support was found (PPE, PPU, _PE and _PU vs. Baseline) (NNE, NNU and _NE vs. Baseline)
▪ Hypothesis 3: $CRP < CRN$	▪ Rejected
▪ Hypothesis 4: $W2WPP = W2WP$; $W2WNN = W2WN$	▪ Support was found (PPE/U vs. _PE/U) (NNE/U vs. _NE/U)
▪ Hypothesis 5: $W2WPN > W2W_N$; $W2WNP < W2W_P$ $W2WPN > W2WNN$; $W2WNP < W2WPP$	▪ Support was found (PNU vs. _NU) (NPU vs. PPU) (PNU vs. NNU)
▪ Hypothesis 6: $CRPP, CRNN > CRPN, CRN$	▪ Support was found (PNE vs. PPE) (PNU vs. PPU) (NPE vs. PPE) (NPU vs. PPU)
▪ Hypothesis 7.1: $W2WPNE < W2WNPE$	▪ Rejected
▪ Hypothesis 7.2 $W2WPNU > W2WNPU$	▪ 7.2 support was found (PNU vs. NPU) (_PE vs. _NE and _PU vs. _NU)
▪ Hypothesis 8: $CRE \neq CRU$	▪ Rejected
▪ Hypothesis 9: $CRPPE - CR_{PE} < CRPPU - CR_{PU}$; $CRNNE - CR_{NE} < CRNNU - CR_{NU}$	▪ Support was found (PPU vs. _PU)
▪ Hypothesis 10: $W2W(P)PE < W2W(N)NE < W2W(N)NU$	▪ Rejected
▪ Hypothesis 11: $W2WPNE < W2WPNU$; $W2WNPE > W2WNPU$	▪ Support was found (PNE vs. PNU)

5. Discussion

The current study provides an investigation about the different effects of the use of different online movie reviews. Analysing the different effects of several review components and the combinations of these review components was the main purpose of this study. Before providing the concluding remarks based on the results from these analyses, it is important to first summarise and discuss these findings.

5.1 Relation to existing findings

The statistical results of this study indicate that there are no different effects between the four main conditions. This means that there is no different effect in willingness to watch a movie and perceived review credibility between positive and negative reviews, congruent and incongruent reviews and expert and user reviews. These findings suggest that it is necessary to examine on a sub condition level to find and map the different effects of online reviews.

First of all, the use or no use of online movie reviews have an influence on the willingness to watch a movie. However, this study indicates that this influence occurs only if the online movie reviews have a negative content and in case it is written by a user it must contain a negative star rating. Therefore, this study underlines the suggestions of Brown & Reingen 1987; Weinberger, Allen & Dillon 1981; Fiske, 1980 and Arndt, 1967. According to the findings of this study, there is no stronger effect occurring in willingness to watch a movie, when the similar review text is supplemented with a matching star rating. In addition, adding a non-matching star rating to a review text written by a user will result in an anchoring effect. The findings underline the expectation on the dominant role of star rating with respect to the review text in case a user is the source. This finding is in the line with previous research findings of Bickart & Schindler, 2001. Last mentioned findings suggested that consumers relate to consumers' experiences easier which may correlate with the approach with which participants interpret the review, by assigning more weight to the star rating.

The reason for the absence of an anchoring effect during interpretation of the expert incongruent reviews could be related to the compensation process of the star rating as well as the review text, which may lead to a levelling effect. Furthermore, in the situation the online reviews contain a review text without a star rating, the expert reviews have more impact compared to

user reviews. In general, the use of expert reviews has a bigger chance to affect the willingness to watch a movie in comparison to the use of user reviews. Therefore, this study underlines previous findings of Vermeulen, Das, & Swager, 2008; Sen & Lerman, 2007 and Senecal & Nantel, 2004. Hereby, is necessary to take the important role of review text into account. In addition, adding a star rating is valuable only in cases where the review is written by a user. In general, the findings of this research correspondent with the existing findings that review designs can influence consumers' processes of interpreting the review (Zhou & Guo, 2017; Agnihotri & Bhattacharya, 2016; Schindler and Bickart, 2012).

The main findings of the current study related to the perceived review credibility were unexpected. Presumably, during the survey a halo-effect was present. However, what makes the presence of halo-effect unlikely, is the survey design. Which refers to the different answer methods of the questions, such as horizontal points and vertical words, and the use of page breaks after each question. In that respect, review users tend to be influenced by their own review assessment of the review content valence.

5.2 Academic implications

Many studies examined the effects of online reviews. However, this is the first study which examines the different effects of star rating in combination with other review components. There is a need for systematic research in this field, especially because of the increased presence of consumer behaviour online. Therefore, the need arises for a comprehensive framework to recognize and to indicate the effects of online consumer behaviour with respect to the use of online reviews. Many existing studies measured the effect of the review frequencies; the valence of the total review; different effect of sources; review styles; the reason to write a review, review helpfulness, review length, review recentness and review use and purchases (Zhou & Guo, 2017; Agnihotri & Bhattacharya, 2016; Jin, Hu & He, 2014; Flanagan & Metzger, 2013; Schindler & Bickart, 2012; Zhao & Xie, 2011; Vermeulen & Seegers 2009; Chen & Xie, 2008; Hu, Liu, & Zhang, 2008; Xia & Bechwati, 2008; Vermeulen, Das, & Swager, 2008; Sen, & Lerman, 2007; Park, Lee, & Han, 2007; Chevalier & Mayzlin, 2006; Ghose, & Ipeirotis, 2006; Senecal, & Nantel, 2004; Chatterjee, 2001; Eliashberg & Shugan, 1997). Despite these different conducted studies, all of these studies indicate there is a need to investigate online reviews and the effects of it. These studies proved that the field of reviews has influence on

consumer behaviour online as well as offline. The results of this study are an extension of this finding because those make clear that depending on the combination of certain review conditions (negative content and in case it is written by a user it must contain a negative star rating) online reviews have impact on the willingness to watch a movie. The results underlined the need for examining the effects of online reviews on the level of sub conditions. This indicates at which level of future academic review research will find relevant outcomes. Furthermore, the results of this study suggest that we have to make a distinction between the interpretation of star rating and text with respect to the different sources. This study could be a reason to investigate the relation of review valence and the perceived review credibility. In addition, the results of this research are based on the assuming that people use one online review and interpret the complete review, the found effects could be a reason to research these effects without this assumption.

5.3. Practical implications

As above mentioned the use of an online collection of information and the online consumer behaviour increased, which leads to a greater need for online reviews. In addition, because of the overload of online information, the online reviews get a key role in considerations of purchases online as well as offline. In case of a website containing product ratings and reviews, people are 63% more likely to purchase a product from a site (Tobin, 2013). Nowadays, 46 % of consumers pay attention and consider the quality of reviews which is increased by 11% with respect to 2016 (BrightLocal, 2017). The review quality became more important in the consumers' minds. Both review platform owners and the review users will benefit from improvements of online reviews. The review platform owners could deliver more appropriate review design, which will lead to a higher platform traffic and a higher amount of website visits as well (BrightLocal, 2017). For companies, research proved that reviews still have a significant impact on product sales (Lu et al, 2013). Besides this, the opportunity to serve more appropriate review design and to promote the existing more appropriate reviews will lead to a more persuasive result. Based on the existing research findings, when companies promote the more persuasive reviews it could influence the purchase intentions in a positive way (Park, etc., 2007). At the same time, review users will interpret the improved reviews easier and better. The use of online reviews could be more accessible, for instance by eliminating the review text of user reviews and or the star rating of expert reviews, which could result in an increase of online

review use. Furthermore, many product/service suppliers and marketers could reach the needs of their consumers better, by analysing the improved reviews easier and better. For instances, to recognize the different role of a star rating of expert or user review. With this, suppliers and marketers can understand the effect of an existing review in a better way. In addition, the results of this study provide that review writers have to be aware that they have only impact on willingness to watch a movie in case they write a review consists of a negative content. Moreover, the found effects indicate that experts have to be aware of the impact of their review text and the levelling effect of their star rating. On the other hand, users have to be aware of the impact of their star rating and the subordinate role of their review text with respect to the willingness to watch a movie.

5.4 Limitations and Future Research

Despite the current results about the different effects of several review components (Star rating, review text and sources) on the willingness to watch a movie and on the perceived review credibility, a few limitations of this study should not remain unmentioned. First, the relatively small scale of this study. The used total sample size is not small (520), however, the different effects were measured for each sub condition (13) separately. Furthermore, during the survey conduction, all participants were located in the United States and most of them had an age between 25 and 44 years. To enhance the insights of this current principal theme, future research has to take a wider range of respondents into account. In addition, this study has the focus on the online review related to movies. To get a complete view of the different effects of several review components, future research has to measure these effects at least on dissimilar industries. Moreover, future research could examine the different influences of these dissimilar industries in relation to the effects of review components.

Studies suggest that during decision making, consumers prefer stimulus-based options, for instance, brands which are observable, over memory-based options, brands which are encountered previously. This occurs especially during ad hoc decision making (Alba, Marmorstein, & Chattopadhyay, 1992). This current study is based on survey outcomes, whereby a fictional movie title and a couple of realistic movie details were shown. With the conducted survey, it was not possible to measure the different effects of the stimulus and memory biased options. Future research could examine this described phenomenon.

Furthermore, many studies indicate that the decision-making process is influenced by unconscious factors (Newell & Shanks, 2014; Dijksterhuis, 2004). This study measures different effects of unconscious factors (star rating, text and sources) on participant's consciousness expressed outcomes (to what extent do you want to watch this movie and do you think the online movie review is credible). Future research could apply online implicit measurement techniques. In contrast to the explicit measures of the current study, implicit measures are less vulnerable related to demand characteristics. Implicit measurement techniques take place outside of consumer awareness (Fazio & Olson, 2003; Banaji, 2001). Further research can implement EEG and Eye Tracking techniques to collect complete insights into the consumer's mind process during the interpretation of online reviews. EEG techniques supply the ability to expose emotional and motivational process in the mind of consumers by means of brain activities (Cooper, Osselton & Shaw, 2014). Applying Eye Tracking techniques make it possible to find the elements which attract the most attention (Poole & Ball, 2006). For instance, in case these techniques are applied while examining the perceived review credibility, it could determine the absent/present of halo-effect during answering the survey or the influences of review valence on the perceived review credibility.

In addition, further research can take response latencies into account. Especially, measuring the response latencies for each question in contrast to current study which compared response latencies of the total survey, whereby interesting results were absent.

Prior research mostly examines the review effects by means of sales figures related to reviews in naturalistic settings (Basuroy, Chatterjee, & Ravid, 2003). This study examines the effects of review components in an experimental setting. Controlled conditions enable measuring the actual effects of review components separately. Although, participants could manipulate their answers consciously or unconsciously to their perceptions of the intention of an experimenter. As a result, future research should examine the effects of the review components by taking the study findings of experimental as well as naturalistic settings into account.

Furthermore, the current study findings are valuable in the stage where online reviews already are selected. Future research can examine effects of the review components related to the review selection. For instances, what is the effect of the review contains only of star rating compared text. In addition, this study is based on the assumptions that people are evaluate one single online review before they make a decision and they interpret this complete review. People could act differently in the reality. This may lead to the limitation of the meaning of the current study

results. When future research takes this into account, it could find the total levelling effect of the current found sub conditions effects. In addition, future research could find the anchoring effect or could find that people do not interpret or read the complete review by applying the EEG and Eye Tracking techniques.

6. Conclusion

The results of this study deliver new insights on the use of online movie reviews and the different effects of several review components. This study suggests that it depends on the combination of review components, star rating, review text and review source, as to what extent an online movie review has an impact on the willingness to watch a movie. It is found that an anchoring effect is only present in specific circumstances of the review components. Furthermore, the star rating, as well as the review text, play different roles depending on whether a review contained a star rating as well as which source is mentioned. In addition, to what extent the contradictory internal review messages is compensated by a consumer depends on the combinations of the review components as well.

The main findings of the perceived review credibility were not directly related to previous expectations. Future research is necessary to assign the findings value.

7. References

- Adelson, J. L., McCoach, D. B., (2010). 'Measuring the Mathematical Attitudes of Elementary Students: The Effects of a 4-Point or 5-Point Likert-Type Scale'. *Educational and Psychological Measurement*, Vol 70, pp. 796 - 807.
- Agnihotri, A., Bhattacharya, S. (2016). 'Online Review Helpfulness: Role of Qualitative Factors'. *Psychology & Marketing*, Vol 33, pp. 1006 - 1017
- Anderson, K.M., Palen, L. and Vieweg, S. (2010), 'Supporting 'Everyday Analysts' in Safety- and Time-Critical Situations', *The Information Society*, Vol 12, pp. 52-62
- Armstrong, Arthur R. and John Hagel III (1996), 'The Real Value of On-Line Communities', *Harvard Business Review*, Vol 74, pp. 134-141.
- Ajzen, I., (1991). 'The Theory of Planned Behavior'. *Organizational, Behavior and Human Decision Processes*, Vol 50, pp. 178 - 211.
- Alba, J., Chattopadhyay, A. (1986). 'Salience Effects in Brand Recall'. *Journal of Marketing Research*, Vol 23, pp. 363–369
- Alba, J., Marmorstein, H., and Chattopadhyay, A. (1992). 'Transitions in preference over time: the effects of memory on message persuasiveness'. *Journal of Marketing Research*, Vol 29, pp. 406–416.
- Allenby, G., Gilbride, T. and Yang, S. (2005), 'Modelling Simultaneity in Survey Data', *Quantitative Marketing and Economics*, Vol 26, pp. 311-335
- Allenby, G., Rossi, P. E. (2003), 'Bayesian Statistics and Marketing', *Marketing Science*, Vol 22, pp. 304-328
- Allsop, D. T., Bassett, B. R. and Hoskins, J. A., (2007). 'Word-of-mouth research: Principles and applications'. *Journal of Advertising Research*, Vol 47, pp. 398.
- Arndt, J., (1967). 'Role of product-related conversations in the diffusion of a new product'. *Journal of Marketing Research*, pp. 291-295.
- Avery, C., Resnick, P. and Zeckhauser, R. (Jun. 1999), 'The market for Evaluations', *The American Economic Review*, Vol 89, pp. 564-584
- Banaji, M. R., Roediger III, H. L., Nairne, J. S., Neath, I., and Surprenant, A. (2001). 'Implicit attitudes can be measured'. *The nature of remembering: Essays in honor of Robert G. Crowder*.
- Bakeman, R. (2005). 'Recommended effect size statistics for repeated measures designs'. *Behavior research methods*, Vol 37, pp. 379-384.

Barnes, S. and Chen, Y. (2007) 'Initial trust and online buyer behaviour', *Industrial Management & Data Systems*, Vol. 107, pp. 21-36

Basuroy, S., S. Chatterjee, and S. A. Ravid (2003, October). 'How critical are critical reviews? the box office effects of film critics, star power, and budgets'. *Journal of Marketing* Vol 67, pp. 103

Bickart, B., Schindler, R. (2001). 'Internet forums as influential sources of consumer information', *Journal of Interactive Marketing*, Vol 15, pp. 31-40.

Bone, P. F., (1995). 'Word-of-mouth Effects on Short-Term and Long-Term Product Judgments', *Journal of Business Research*, Vol 32, pp. 213 - 223

Brown, J. J., Reingen P. H. (1987), 'Social Ties and Word-of Mouth Referral Behaviour', *Journal of Consumer Research*, pp. 350-362.

BrightLocal, (2017), 'Local Consumer Review Survey', *BrightLocal*. [Online] Available at: <https://www.brightlocal.com/learn/local-consumer-review-survey/> [Accessed 18 December 2017].

Çelebi, S. I. (Oct. 2011), 'The effect of the mass media and demographics on pre-purchase, purchase and post-purchase activities', *Analysis and Metaphysics*, Vol 10, pp. 67-80

Chatterjee, P. (2001), 'Online reviews: Do consumers use them?', *Advances in Consumer Research*, Vol 28, pp. 129-133

Chen, Y., Xie, J. (Mar. 2008), 'Online consumer review: Word-of-Mouth as a New element of marketing communication mix', *Management Science*, Vol 54, pp. 477-491

Chevalier, Judith A., and Dina Mayzlin (2006), 'The effect of word of mouth on sales: Online book reviews.', *Journal of marketing research*, Vol 43 pp. 345-354

Cheung, C. M., Thadani, D. R., (2012). 'The Impact of Electronic Word-of-Mouth Communication: A literature Analysis and Integrative Model'. *Decision Support Systems*, Vol 54, pp. 461 - 470.

Cohen, J. (1992). 'A power primer'. *Psychological bulletin*, Vol 112, pp. 155.

Cooper, R., Osselton, J. W., & Shaw, J. C. (2014). *EEG technology*. Butterworth-Heinemann.

Court, D., Elzinga, D., Mulder, S. and Vetvik, O. J., (2009). 'The consumer decision journey' *McKinsey*. [Online] Available at: <https://www.mckinsey.com/business-functions/marketing-and-sales/our-insights/the-consumer-decision-journey> [Accessed 16 December 2017].

Dam, van J, (2006) 'Bij een goed cijfer is het effect groot', *Het Parool*

Dijksterhuis, A. (2004). 'Think different: the merits of unconscious thought in preference development and decision making'. *Journal of personality and social psychology*, Vol 87, pp. 586.

Dimensional Research (2013), 'Customer service and business results: A survey of customer service from mid-size companies', *Dimensional Research*, Vol 11

Duan, W., Gu, B. and Whinston, A. B., (2008). 'Do online reviews matter? - An empirical investigation of panel data'. *Decision Support Systems*, pp. 1007 - 1016.

East, R., Hammond, K. and Lomax, W. (Feb. 2008), 'Measuring the impact of positive and negative word of mouth on brand purchase probability', *International journal of research in marketing*, Vol 25, pp. 215-224

Eliashberg, J., Shugan, S.M. (1997), 'Film Critics: Influencers or Predictors?', *Journal of Marketing*, Vol 61, pp. 68-78

Epley, N., and Gilovich, T. (2006). 'The anchoring-and-adjustment heuristic: Why the adjustments are insufficient'. *Psychological science*, Vol 17, pp. 311-318.

Fazio, R., Olson, M. (2003). 'Implicit measures in social cognition research: their meaning and use'. *Annual Review of Psychology*, pp. 297-328.

Ferguson, C. J. (2009). 'An effect size primer: A guide for clinicians and researchers'. *Professional Psychology: Research and Practice*, Vol 40, pp. 532.

Fishbein, M., Ajzen, I., (1975). 'Belief, Attitude, Intention and Behavior: An Introduction to Theory and Research'. *United States of America: Addison-Wesley Publishing Company*.

Fiske, S. T., (1980). 'Attention and Weight in Person Perception: The Impact of Negative and Extreme Behavior'. *Journal of Personality and Social Psychology*, Vol 38, pp. 889 - 906.

Flanagin, A. J., Metzger, M. J. (2008), 'Digital media and youth: Unparalleled opportunity and unprecedented responsibility', *Digital media, youth, and credibility*, Vol 201, pp. 5-27

Flanagin, A. J., Metzger, M. J. (2013), 'Trusting expert- versus user-generated ratings online: The role of information volume, valence, and consumer characteristics', *Computers in Human Behavior*, Vol 8, pp. 1626-1634

Garbarino, E., Johnson, M.S. (1999) 'The different roles of satisfaction, trust, and commitment in customer relationships,' *Journal of Marketing*, Vol 63, pp. 70-87.

Gefen, D. (2002). 'Customer loyalty in e-commerce'. *Journal of the association for information systems*, Vol 3, pp.27-51

Gefen, D., Karahanna, E., and Straub, D.W., (2003), 'Trust and TAM in Online Shopping: An Integrated Model,' *MIS Quarterly* Vol 27, pp. 51-90.

Gefen, D. Karahanna, E., and Straub, D.W., (2003) 'Inexperience and Experience with online stores: The importance of TAM and Trust', *IEEE Transactions on Engineering Management*, Vol 50, pp. 307-321.

Ghose, A., Ipeirotis, P. G. (2006), 'Designing systems for consumer reviews: The impact of review subjectivity on product sales and review quality', *Department of information, Operations and Management Sciences*

Ghosh, D., Vogt, A. (2012). 'Outliers: An evaluation of methodologies', *Joint Statistical Meetings*, pp. 3455-3460. San Diego, CA: American Statistical Association

Gino, F., Moore, D.A. (2007), 'Effects of task difficulty on use of advice', *Journal of Behavioral Decision Making*, Vol 20, pp. 21-35

Hao, Y. Y., Ye, Q., Li, Y. and Cheng, Z. J., (2010). 'How Does the Valence of Online Consumer Reviews Matter in Consumer Decision Making? Differences between Search Goods and Experience Goods'. *IEEE Computer Society*, pp. 1 - 10

Hennig-Thurau, T., Gwinner, K., Walsh, G., and Gremler, D. (2004). 'Electronic word-of-mouth via consumer-opinion platforms: what motivates consumers to articulate themselves on the Internet?' *Journal of Interactive Marketing*, Vol 18, pp. 38-52.

Hoch, S. J., Ha, Y.-W., (1986). 'Consumer Learning: Advertising and the Ambiguity of Product Experience'. *Journal of Consumer Research*, Vol 13, pp. 221-233.

Hoeffler, S., Ariely, D. (1999). 'Constructing stable preferences: a look into dimensions of experience and their impact on preference stability'. *Journal of Consumer Psychology*, Vol 8, pp. 113-139.

Hoffman, D.L., Novak, T.P., and Peralta, M.,(1999) 'Building Consumer Trust Online', *Communications of the ACM*, Vol 42, , pp. 80-85.

Hong, W., Thong, J.Y.L., and Tam, K.Y. (2004), 'The Effects of Information Format and Shopping Task on Consumers' Online Shopping Behaviour: A Cognitive Fit Perspective'. *Journal of Management Information Systems*, Vol. 21, pp. 149-184

Hu, N., Liu, L., and Zhang, J. J. (2008), 'Do online reviews affect product sales? The role of reviewer characteristics and temporal effects'. *Information Technology and Management*, Vol 9, pp. 201-214

Huang, J.-H., Chen, Y.-F., (2006). 'Herding in Online Product Choice', *Psychology & Marketing*, Vol 23, pp. 413 - 428.

Huang, P., Lurie, N. H. and Mitra, S., (2009). 'Search for Experience on the Web: An Empirical Examination of Consumer Behavior for Search and Experience Goods'. *Journal of Marketing*, Vol 73, pp. 55-69.

Janssens, W., De Pelsmacker, P., and Van Kenhove, P., (2008). *Marketing research with SPSS*. Pearson Education.

Jarvenpaa, S.L., Tractinsky, N., and Vitale, M., (2000) 'Consumer trust in an Internet store,' *Information Technology and Management*, pp. 45-71.

Jin, L., Hu, B. and He, Y., (2014). 'The Recent versus The Out-Dated: An Experimental Examination of the Time-Variant Effects of Online Consumer Reviews'. *Journal of Retailing*, Vol 90, pp. 552 - 566

Jindal, N., Liu, Y. (2008), 'Word of Mouth for Movies: Its Dynamics and Impact on Box Office Revenue', *Journal of Marketing*, Vol 70, pp. 74–89

Kanouse, D.E. (1984), 'Explaining negativity biases in evaluation and choice behavior: Theory and research', *Advances in Consumer Research*, pp. 703–708.

Khan, A., Rayner, G. D. (2003). 'Robustness to Non-Normality of Common Tests for the Many- Sample Location Problem'. *Journal of Applied Mathematics and Decision Sciences*, Vol 7, pp. 187-206.

Kietzmann, J., Canhoto, A., (2013). 'Bittersweet! Understanding and Managing Electronic Word of Mouth'. *Journal of Public Affairs*, Vol 13, pp. 146-159.

Klucharev, V., Schmidts, A. and Fernandez, A. (2008), 'Brain Mechanisms of Persuasion: How "Expert" Power Modulates Memory and Attitudes', *Social Cognitive and Affective Neuroscience*, Vol 3, pp. 353–366.

Lamb, J. (2018), Cover page, mage code: 161098250, retrieved from <https://www.gettyimages.nl/license/161098250>

List, J. A., Sadoff, S. and Wagner, M., (2011). 'So You Want to Run an Experiment, Now What? Some Simple Rules of Thumb for Optimal Experimental Design?'. *Exp. Econ.* Vol 14, pp. 439 - 457.

Lynch, J. G., Ariely, D., (2000). 'Wine Online: Search Costs Affect Competition on Price, Quality and Distribution'. *Marketing Science*, Vol 19, pp. 83-103.

Mizerski, R. W. (1982), 'An attribution explanation of the disproportionate influence of unfavourable information', *Journal of Consumer Research*, Vol 9, pp. 301-310

Morrison, D. G., (1979). 'Purchase Intentions and Purchase Behavior'. *Journal of Marketing*, Vol 43, pp. 65 - 74.

Morwitz, V., (2012). 'Consumers' Purchase Intentions and their Behavior'. *Foundations and Trends in Marketing*, Vol 7, pp. 181 - 230.

Mudambi, S. M., Schuff, D. (2010). 'What Makes a Helpful Online Review? A Study of Customer Reviews on Amazon.com'. *MIS Quarterly*, Vol 34 , pp. 185-200.

Nash Information Services, The Numbers, 'Domestic Movie Theatrical Market Summary 1995 to 2017', website <http://www.the-numbers.com/market/>, accessed January 2017

Nedungadi, P. (1990). 'Recall and consumer consideration sets: influencing choice without altering brand evaluations'. *Journal of Consumer Research*, Vol 17, pp. 263–276.

Nelson, P. (1970). 'Information and consumer behavior'. *Journal of political economy*, Vol 78, pp. 311-329.

- Nelson, P., (1974) 'Advertising as Information'. *Journal of Political Economy*, p. 730
- Newell, B. R., Shanks, D. R. (2014). 'Unconscious influences on decision making: A critical review'. *Behavioral and Brain Sciences*, Vol 37, pp. 1-19.
- Nielsen, (2015), 'Global trust in advertising', *Nielsen*, Vol 22, pp. 4-8 [Online], Available at: <https://www.nielsen.com/content/dam/nielsen-global/apac/docs/reports/2015/nielsen-global-trust-in-advertising-report-september-2015.pdf>. [Accessed 1 November 2017]
- Park, C., Lee, T. M., (2009), 'Information Direction, Website Reputation and EWom Effect: A Moderating Role of Product Type'. *Journal of Business Research*, Vo 62, pp. 61 – 67
- Park, D.H., Lee, J., and Han, I. (2007), 'The Effect of On-Line Consumer Reviews on Consumer Purchasing Intention: The Moderating Role of Involvement', *International Journal of Electronic Commerce*, Vol. 11, pp. 125-148
- Pearl, J. (1978). 'On the connection between the complexity and credibility of inferred models'. *International Journal of General System*, Vol 4, pp. 255-264.
- Purnawirawan, N., De Pelsmacker, P. and Dens, N., (2012). 'Balance and Sequence in Online Reviews: How Perceived Usefulness Affects Attitudes and Intentions.' *Journal of Interactive Marketing*, Vol 25, pp. 244-255.
- Petty, R.E., Cacioppo, J.T.(1986), 'From Communication and Persuasion: Central and Peripheral Routes to Attitude Change.' New York: Springer-Verlag.
- Philips, D. M., Olsen, J. C. and Baumgartner, H., (1995). 'Consumption Visions in Consumer Decision Making'. *Advances in Consumer Research*, Vol 22, pp. 280 - 284.
- Poole, A., Ball, L. J. (2006). 'Eye tracking in HCI and usability research'. *Encyclopedia of human computer interaction*,
- Priester, J., Nayakankuppam, D., Fleming, M., and Godek, J. (2004). The ASC model: the influence of attitudes and attitude strength on consideration and choice. *Journal of Consumer Research*, 30(4), 574–587.
- Ramaswamy, S., (2013). 'Shopping Then and Now: Five Ways Retail Has Changed and How Businesses Can Adapt', *Think With Google*. [Online] Available at: <https://www.thinkwithgoogle.com/advertising-channels/mobile/five-ways-retail-has-changed-and-how-businesses-can-adapt> [Accessed 19 December 2016]
- Reinstein, D. A., C. M. Snyder (2000, September). 'The influence of expert reviews on consumer demand for experience goods: A case study of movie critics'. *Working paper, George Washington University*.
- Schindler, R. M., Bickart, B.(Mar. 2012), 'Perceived helpfulness of online consumer reviews: The role of message content and style', *Journal of consumer behaviour*, Vol 11, pp. 234-243

Sen, S., Lerman, D. (2007), 'Why are you telling me this? An examination into negative consumer reviews on the web', *Journal of Interactive Marketing*, Vol 21, pp. 76-94

Senecal, S., Nantel, J. (2004). 'The influence of online product recommendations on consumers' online choices', *Journal of Retailing*, Vol 80, pp. 159–169.

Shrestha, K., (2016) '50 Stats you need to know about online Reviews', *Vendasta*, [Online], Available at: <https://www.vendasta.com/blog/50-stats-you-need-to-know-about-online-reviews> [Accessed 18 December 2017].

Simon, H. A. (2013). 'Administrative behavior'. *New York: Simon and Schuster*, 4th ed.

Statista, 'Global online TV and video revenue from 2010 to 2020, by source (in billion U.S. dollars)', website <https://www.statista.com/statistics/246580/global-online-tv-and-video-revenue-by-source/>, Updated 2017, accessed January 2017

Skowronski, J. J., Carlston, D. E. (1987), 'Social Judgment and Social Memory: The Role of Cue Diagnosticity in Negativity, Positivity, and Extremity Biases'. *Journal of Personality and Social Psychology*, vol 52, pp. 689-699.

Sorensen, A. T., Rasmussen, S. J. (2004). 'Is any publicity good publicity? A note on the impact of book reviews'. *NBER Working paper, Stanford University*.

Sparks, B. A., Browning, V., (2011). 'The impact of online reviews on hotel booking intentions and perception of trust'. *Tourism Management*, Vol 32, pp. 1310 - 1323

Thorson, K., Vraga, E., and Ekdale, B. (2010). 'Credibility in context: How uncivil online commentary affects news credibility'. *Mass Communication and Society*, Vol 13, pp. 289-313.

Tobin, S., (2013), 'Why online reviews matters', *SocialMediaToday*, [Online], Available at: <https://www.socialmediatoday.com/content/why-online-reviews-matter> [Accessed 18 December 2017].

Tversky, A., Kahneman, D. (1974). 'Judgment under uncertainty: Heuristics and biases.', *Science*, vol 185, pp. 1124–1130

Vermeulen, I. E., Seegers, D. (2009). 'Tried and tested: The impact of online hotel reviews on consumer consideration', *Tourism management*, Vol 30, pp. 123-127.

Vermeulen, I., Das, E., and Swager, M. (2008). 'Strong medicine or bitter pill? The impact of online drug reviews on drug purchase'. In F. Costa Pereira, J. Verissimo, & P. Neijens (Eds.), *New trends in advertising research*. Lisbon: Edição Si'labo.

Wang, Z. et al., (2012). 'Social Media Recommendation'. *Social Media Retrieval*. s.l.:s.n., pp. 23-42.

Welke, B. (2015), 'Online Reputation and online Reviews Statistics You Need to Know', *RenegadeWorks*, [Online], Available at: <https://renegadeworks.com/online-reputation-reviews-statistics/> [Accessed 18 December 2017].

Weinberger, M.G., Allen, C.T., and Dillon, W.R. (1981), 'Negative information: Perspectives and research directions'. *Advances in Consumer Research*, Vol 8, pp. 398- 404

Whitlark, D. B., Geurts, M. D., and Swenson, M. J. (1993). 'New product forecasting with a purchase intention survey'. *The Journal of Business Forecasting Methods & Systems*, Vol 12, pp. 18.

Wüthrich, S. (2015), 'Infographic: Everything you need to know about online reviews', *Web republic*, [Online], Available at: <https://webrepublic.com/en/blog/2015/1/29/infographic-everything-about-online-reviews> [Accessed 18 December 2017].

Xia, L., Bechwati, N. N. (2008), 'Word of mouse: The role of cognitive personalization in online consumer reviews', *Journal of Interactive Advertising*, Vol 9, pp. 3-13

Yakov, B., Shankar V., Sultan, F. and Urban, G. L. (2005), 'Are Drivers and Role of Online Trust the Same for All Web Sites and Consumers' *Journal of Marketing*, Vol 69, pp. 133-152

Zhao, M., Xie, J. (Jun. 2011), 'Effects of social and Temporal distance on consumers' responses to peer recommendations', *Journal of Marketing research*, Vol 48, pp. 486-496

Zhou, S., Guo, B. (2017). 'The order effect on online review helpfulness: A social influence perspective'. *Decision Support Systems*, Vol 93, pp. 77 - 87

Appendix 1 – Pre-test

Thank you for your participation in this pre-test survey of my master's thesis.

This survey consists of a number of questions. The results of this survey will be used as a base of the final survey.

All responses will be processed anonymously and not made available to third parties.

Please read the following text about a movie

Movie title: Princess and the dragon

'If you want to pick holes, the Dragon is a bit of drip. And while the Princess is a young woman in charge of her destiny, it's not totally clear what role there is for her beyond wifely fey for lively after her happy ending. Did the world need another 'Princess and Dragon? Maybe not, but this one is still a keeper'.

Source: www.metacritic.com

How positive or negative is the text?

Extremely negative									Extremely positive	
0	1	2	3	4	5	6	7	8	9	10
<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Please read the following text about a movie

Movie title: Princess and the dragon

'The actress is a perfect princess, but ultimately it doesn't matter. This film is what it is. It's not a bad movie in most respects, there is plenty of good craftsmanship on display, but it isn't art and it doesn't aspire to be. It's what a cheeseburger would taste like in purgatory'.

Source: www.metacritic.com

How positive or negative is the text?

Extremely negative										Extremely positive
0	1	2	3	4	5	6	7	8	9	10
<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Please read the following text about a movie

Movie title: Escape

'It is good to see a movie go back to being what movies actually are, motion with dialogue. Most movies forget about the motion and just focus on dialogue. It was a blast and left me cheering and satisfied at the end. The screenplay was cleverly written to scare the crap out of you suddenly and at different scenes interject street humour. A few racial uncomfortable overtones mixed in. Sometimes, it was hard to follow the dialogues and the situations. It made me restless during the movie'.

Source: www.metacritic.com

How positive or negative is the text?

Extremely negative										Extremely positive
0	1	2	3	4	5	6	7	8	9	10
<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Please read the following text about a movie

Movie title: Escape

'I think that this film was very strange. At some points, I thought to myself "Am I watching a horror movie or a thriller?". The cast did an okay job with the film. The plot line was very similar to A Cure For Wellness. The part with the deer made out of bones was way too weird for my taste'.

Source: www.metacritic.com

How positive or negative is the text?

Extremely negative										Extremely positive	
0	1	2	3	4	5	6	7	8	9	10	
<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	

Please read the following text about a movie

Movie title: Isolation

'I thought it was a really good movie. It was realistic and it was really cool. The action sequences were cool like in the first twenty minutes and it got better from there. It was really cool to see a group of American soldiers overcome a stronger German army in a cool way. After the first twenty minutes, the movie became somewhat boring and stereotyped.'

Source: www.metacritic.com

How positive or negative is the text?

Extremely negative										Extremely positive	
0	1	2	3	4	5	6	7	8	9	10	
<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	

Please read the following text about a movie

Movie title: Isolation

'Some interesting direction in the opening and climactic battle scenes, but the film is let down by the plot, which is too simple and implausible to engage the viewer. It's a long and cliched Hollywood blockbuster war film and not much more. During the film you really believe the actors and you will not get distracted.'

Source: www.metacritic.com

How positive or negative is the text?

Extremely negative										Extremely positive	
0	1	2	3	4	5	6	7	8	9	10	
<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	

Appendix 2 – Pre-test Outcomes

Table 4
PRE-TEST OUTCOMES

Review	Princess and the Dragon - 1	Average score: 7.1
Review	Princess and the Dragon - 2	Average score: 3.4
Review	Escape - 1	Average score: 7.0
Review	Escape - 2	Average score: 3.3
Review	Isolation - 1	Average score: 7.7
Review	Isolation - 2	Average score: 5.3

Note: The sample size was 15.

Appendix 3 - Final Survey

Q1 Introduction

You are invited to participate in this study for a master thesis at Erasmus University, the Netherlands. Before you start to answer the questions, I kindly would like to ask you to read this introduction carefully.

The intention of this study is to research and collect knowledge of online consumer behaviour. Answering the questions will approximately take 3 minutes in total. If you at any point wish to stop the study, you have the right to do and there will be no consequences. Your data will be anonymous. There will be no record that links the collected data based on your given answers with any personal data from which you could be identified (e.g. demographics, contact related data).

Q 2.0 Please read the following movie information.

Title: Princess and the dragon

Summary: A bright, beautiful and independent young woman is taken prisoner by a dragon in his castle. Despite her fears, she befriends the castle's enchanted staff and learns to look beyond the Dragon's hideous exterior and realise the kind heart and soul of the true person within.

Genre(s): Fantasy, Romance, Family, Musical

Runtime: 129 min

Source: www.metacritic.com

Q 2.1 Please read the following online movie review.

Lucy (Movie Expert)

'The actress is a perfect princess, but ultimately it doesn't matter. This film is what it is. It's not a bad movie in most respects, there is plenty of good craftsmanship on display, but it isn't art and it doesn't aspire to be. It's what a cheeseburger would taste like in purgatory.'



Q 2.2 Please express to what extent you want to watch this movie.

Definitely not					Neutral		Definitely yes			
0	1	2	3	4	5	6	7	8	9	10

Q 2.3 How credible was this online movie review?

Not at all credible
Not credible
Somewhat credible
Credible
Very credible

Q 2.4 Please read the following movie information.

Title: Escape

Summary: Now that James and his girlfriend, Jennifer, have reached the meet-the-parents milestone of dating, she invites him for a weekend getaway upstate with Kate and Daniel. At first, James reads the family's overly accommodating behaviour as nervous attempts to deal with their daughter's interracial relationship, but as the weekend progresses, a series of increasingly disturbing discoveries lead him to a truth that he could have never imagined.

Genre(s): Mystery, Thriller, Horror, Comedy

Runtime: 104 min

Source: www.metacritic.com

Q 2.5 Please read the following online movie review.

Chris (Movie Expert)

'I think that this film was very strange. At some points, I thought to myself "Am I watching a horror movie or a thriller?". The cast did an okay job with the film. The plot line was very similar to A Cure For Wellness. The part with the deer made out of bones was way too weird for my taste.'



Q 2.6 Please express to what extent you want to watch this movie.

Definitely not

Neutral

Definitely yes

0	1	2	3	4	5	6	7	8	9	10
---	---	---	---	---	---	---	---	---	---	----

Q 2.7 How credible was this online movie review?

Not at all credible

Not credible

Somewhat credible

Credible

Very credible

Q 3.0 To what extent do you think online movie reviews are helpful by making a decision?

Very helpful

helpful

somewhat helpful

unhelpful

Very unhelpful

Q 3.1 How often do you use online movie reviews?

Never

Sometimes

About half the time

Often

Very often

Q 3.2 Please rank the following movie genres based on your preference, when 1 is the most preferred and 5 is the least preferred

- 1 Comedy
 - 2 Drama
 - 3 Action
 - 4 Art House
 - 5 Thriller
 - 6 Animation
-

Q 3.3 What is your gender?

Male

Female

Q 3.4 What is your age?

Under 18

18 - 24

25 - 34

35 - 44

45 - 54

55 - 64

65 - 74

75 - 84

85 or older

Appendix 4 - Statistical Outcomes

Table 5.1
WILLINGNESS
POSITIVE REVIEW AND BASELINE

One simple t-test		
	Sig. (2-tailed)	Mean Difference
Positive review	,000	6,08176
Baseline	,000	5,84783

Table 5.2
WILLINGNESS
NEGATIVE REVIEW AND BASELINE

One simple t-test		
	Sig. (2-tailed)	Mean Difference
Negative review	,000	4,66327
Baseline	,000	5,84783

Table 5.3
WILLINGNESS
POSITIVE REVIEW AND NEGATIVE REVIEW

One simple t-test		
	Sig. (2-tailed)	Mean Difference
Positive review	,000	6,08176
Negative review	,000	4,66327

Table 5.4
WILLINGNESS
REVIEW SOURCE

One simple t-test		
	Sig. (2-tailed)	Mean Difference
Expert review	,000	5,30242
User review	,000	5,46239

WILLINGNESS
REVIEW SOURCE AND REVIEW VALENCE

One simple t-test		
	Sig. (2-tailed)	Mean Difference
Expert positive review	,000	5,98276
User positive review	,000	6,20139

Matrix: Willingness to watch a movie

Table 5.5
WILLINGNESS AND TOTAL SUB CONDITIONS

Mutiple Comparisons LSD													
	Baseli ne	In- PNE	In- PNU	In- NPE	In- NPU	Co- PPE	Co- PPU	Co- NNE	Co- NNU	Ot- PE	Ot- PU	Ot- NE	Ot- NU
Base		0,061	0,758	0,347	0,114	0,733	0,143	0,039	0,001	0,823	0,982	0,008	0,104
PNE	0,8923*		0,036	0,369	0,793	0,029	0,002	0,767	0,144	0,038	0,069	0,39	0,848
PNU	-0,1522	-1,0444*		0,229	0,069	0,981	0,258	0,023	0,001	0,927	0,783	0,004	0,063
NPE	0,455	-0,4373	0,6071		0,531	0,208	0,022	0,254	0,024	0,251	0,355	0,088	0,493
NPU	0,7645	-0,1278	0,9167*	0,3095		0,059	0,004	0,589	0,093	0,075	0,124	0,272	0,947
PPE	-0,1638	-1,0561*	-0,0116	-0,6188	-0,9283*		0,257	0,019	0,001	0,906	0,761	0,003	0,054
PPU	-0,7582	-1,6505*	-0,6061	-1,2132*	-1,5227*	-0,5944		0,001	0	0,212	0,164	0	0,004
NNE	1,0423*	0,15	1,1944*	0,5873	0,2778	1,2061*	-1,8005*		0,265	0,024	0,045	0,596	0,639
NNU	1,6509*	0,7586	1,8030*	1,1959*	0,8864*	1,8147*	2,4091*	0,6086		0,001	0,002	0,541	0,11
_PE	-0,1067	-,9990*	0,0455	-0,5617	-0,8712*	0,0571	0,6515	-1,1490*	-1,7576*		0,848	0,005	0,068
_PU	-0,0111	-0,9034*	0,141	-0,4661	-0,7756	0,1527	0,7471	-1,0534*	-1,6620*	0,0956		0,01	0,113
_NE	1,3215*	0,4292	1,4737*	0,8665*	0,557	1,4853*	2,0797*	0,2792	-0,3293	1,4282*	1,3327*		0,308
_NU	0,7978	-0,0944	0,95*	0,3429	0,0333	0,9616*	1,5561*	-0,2444	-0,853	0,9045*	0,809	-0,5237	

Note: Mean Difference (Horizontal title – Vertical title) | Left Bottom Triangle
Sig. | Right Top Triangle

*. The mean difference is significant at the ,1 level.

Baseline = No review;

In- = Incongruent; PNE = Positive star rating, Negative text and Expert; PNU = Positive star rating, Negative text and User; NPE = Negative star rating, Positive text and Expert; NPU = Negative star rating, Positive text and User.

Co- = Congruent; PPE = Positive star rating, Positive text and Expert; PPU = Positive star rating, Positive text and User; NNE = Negative star rating, Negative text and Expert; NNU = Negative star rating, Negative text and User.

Ot- = Only text; _PE = No star rating, Positive text and Expert; _PU = No star rating, Positive text and User; _NE = No star rating, Negative text and Expert; _NU = No star rating, Negative text and User.

Table 6.1
CREDIBILITY
POSITIVE REVIEW AND NEGATIVE REVIEW

One simple t-test		
	Sig. (2-tailed)	Mean Difference
Positive review	,000	3,54403
Negative review	,000	3,07483

Table 6.2
CREDIBILITY
CONGRUENT AND INCONGRUENT REVIEW

One simple t-test		
	Sig. (2-tailed)	Mean Difference
Congruent review	,000	3,29310
Incongruent review	,000	3,26488

Table 6.3
CREDIBILITY
EXPERT REVIEW AND USER REVIEW

One simple t-test		
	Sig. (2-tailed)	Mean Difference
Expert review	,000	3,27218
User review	,000	3,32965

Matrix: Review credibility

Table 6.4

CREDIBILITY AND TOTAL SUB CONDITIONS

Mutiple Comparisons LSD												
	In- PNE	In- PNU	In- NPE	In- NPU	Co- PPE	Co- PPU	Co- NNE	Co- NNU	Ot- _PE	Ot- _PU	Ot- _NE	Ot- _NU
PNE		0,641	0,362	0,855	0,031	0,004	0,007	0,012	0,255	0,419	0,436	0,874
PNU	-,0701		0,181	0,525	0,106	0,019	0,002	0,004	0,527	0,741	0,23	0,763
NPE	0,1341	0,2042		0,474	0,003	0	0,073	0,103	0,043	0,095	0,916	0,298
NPU	0,027	0,0971	-,1071		0,022	0,003	0,013	0,022	0,194	0,331	0,554	0,739
PPE	-,3158*	-,2457	-,4499*	-,3427*		0,386	0	0	0,308	0,201	0,005	0,053
PPU	-,4535*	-,3834*	-,5877*	-,4805*	-,1378		0	0	0,069	0,041	0,001	0,008
NNE	0,4139*	0,4840*	0,2798*	0,3869*	0,7297*	0,8674*		0,909	0	0,001	0,064	0,006
NNU	0,3949*	0,4650*	0,2608	0,3680*	0,7107*	0,8485*	-,0189		0	0,002	0,09	0,01
_PE	-,1657	-,0956	-,2998*	-,1926	0,1501	0,2879*	-,5795*	-,5606*		0,769	0,063	0,344
_PU	-,1214	-,0513	-,2555*	-,1484	0,1944	0,3322*	-,5353*	-,5163*	0,0443		0,127	0,527
_NE	0,1178	0,1879	-,0163	0,0909	0,4336*	0,5714*	-,2961*	-,2771*	0,2835*	0,2392		0,363
_NU	-,0236	0,0465	-,1577	-,0506	0,2922	0,4299*	-,4375*	-,4186*	0,142	0,0978	-,1414	

Note: Mean Difference (Horizontal title – Vertical title) | Left Bottom Triangle

Sig. | Right Top Triangle

*. The mean difference is significant at the ,1 level.

In- = Incongruent; PNE = Positive star rating, Negative text and Expert; PNU = Positive star rating, Negative text and User; NPE = Negative star rating, Positive text and Expert; NPU = Negative star rating, Positive text and User.

Co- = Congruent; PPE = Positive star rating, Positive text and Expert; PPU = Positive star rating, Positive text and User; NNE = Negative star rating, Negative text and Expert; NNU = Negative star rating, Negative text and User.

Ot- = Only text; _PE = No star rating, Positive text and Expert; _PU = No star rating, Positive text and User; _NE = No star rating, Negative text and Expert; _NU = No star rating, Negative text and User.