HOW TO DESIGN A GOOD REPUTATION SYSTEM FOR AN ONLINE PEER-TO-PEER PLATFORM

F.L.M. Veldhuizen
s1314335

Faculty of Electrical Engineering, Mathematics and Computer Science

Human Media Interaction (HMI)

Master Thesis Interaction Technology

supervisor: dr.ir. D. Reidsma
supervisor: dr. J.G. Meijerink
supervisor: dr. M. Theune

UNIVERSITY OF TWENTE.

30TH JANUARY 2020

Abstract

In this thesis, the fundamentals of online reputation systems were explored and explained. After performing a background research, the metrics that can effectively evaluate the performance of a reputation system were defined. The performance of a reputation system is determined by its credibility, along with the intention to review that it elicits. Based on this principle, a conceptual framework was designed that can aid in the creation and evaluation of online reputation systems. The conceptual framework predicts the performance of reputation systems by classifying reputation system features in either 'credibility' or 'intention to review'. This framework was later improved by testing its features in two experiments. The first experiment was exploratory: the goal was to determine the effectiveness of the gathered methods to positively influence credibility and intention to review, and to find out how people perceived these methods. The results of the first experiment led to the second experiment, which focused on four topics: the distinction in behavior between platform users and non-platform users, the degree of which the credibility of a review is influenced by its textual content, reluctance in online user feedback, and finally, how gamification features are perceived by users. Results of both of these experiments led to an improvement in the initial conceptual framework, which concludes the thesis.

Contents

Foreword

Before I explain my personal motivation to write this thesis, I want to thank my supervisors Dennis Reidsma, Jeroen Meijerink and Mariët Theune for their time, expertise and guidance throughout this project. I also want to thank my family and friends, who have shared their insights and have supported me while making the thesis. Finally, I want to thank the Picobelly team: Wesley Lam, Max Meijer, Wouter Kaag, Maksym Aleksandrovych and Thomas Cavas for inspiring me to write this thesis, and for teaching me a lot about myself.

As for my personal motivation, two years ago I founded my own start-up company: Picobelly[1]. Picobelly is a gig economy platform, which is a platform where people do freelance gigs (sellers, in Picobelly's case these are homecooks who cook and serve different meals), that other people (buyers) can pay for. A well known gig economy platform is Uber, where people can offer a taxi service from their own car to generate an extra income.

During the last two years' development of Picobelly, I learned a lot about starting a business. It was a great learning experience with amazing moments, like Wesley and me using our own platform for the first time and having a memorable dinner at an international student's house. There were also problematic moments as starting a business is not an easy task. Many difficulties related to developing an online platform had to be overcome by the team. In the first chapter of the thesis I will elaborate on my motivation to perform a research on one of these vital problems: reputation systems.

---

[1] website: picobelly.com

1 Introduction

1.1 Motivation

One of the most important problems that online platforms struggle with is building a healthy and active community, a community where there is mutual respect and trust among users. Generating mutual trust within an online community (especially on peer-to-peer platforms where users mainly interact with other users) is vital for the success of a platform, as users need to have a realistic and positive expectation of the interaction before engaging with other users (Hawlitschek et al., 2018; Mazzella et al., 2016). There are different ways to generate more trust among users on a platform, but literature shows us that the establishment of an effective reputation system plays a key role in building trust (Thierer et al., 2015). Additionally, reputation systems allow a platform to be autonomously regulated by its own users (to a degree), which makes online platforms very scalable by nature.

However, there are multiple problems when looking at reputation systems. For instance, when products online receive no user feedback or when people lie when writing their reviews (and other users are aware of this), people become sceptical and hesitate before engaging with other users on the platform. By stimulating users to leave honest reviews, trust in a platform and therefore the performance of a platform can be enhanced. More information about problems related to online user reviews can be found in section 2.2.2. As problems involving trust frequently occur on online platforms, it seems that current solutions are not always enough. This thesis aims at making an improvement on the current selection of available reputation systems for online platforms.

1.2 Goal and approach

The goal of this thesis is to find out how online reputation systems work and how a good reputation system can be designed for an online peer-to-peer platform. When this goal is achieved, it will be easier to design high quality reputation systems for online environments in the future. To achieve this goal, it is important to clearly define what makes a reputation system good, and how the performance of a reputation system can be evaluated.

To find out what makes a reputation system effective, a research on existing literature about online platforms and reputation systems was performed. Along with scientific literature, current state-of-the-art solutions were researched and explored. This background research led to the design of a conceptual framework that can be used to identify reputation system features, and to estimate reputation system performance.

The conceptual framework that was designed based on the context analysis cannot be used to evaluate reputation systems effectively without validating its features. To expand and validate this framework (and make it more usable in reputation system design), experiments were performed that explore the reputation system features found in the framework. By performing experiments with different features in reputation systems, an insight into how a reputation system can be optimally designed at a feature level is obtained. By finding out what the impact of different features - and combinations thereof - is on user feedback, hypothetically an effective reputation system that yields better reviews than currently possible can be designed.

1.3 Research Questions

In this section, the research questions that were devised to achieve the thesis goal are presented. The main research question is presented below, and is then divided in multiple sub-questions. These questions together comprise the different stages of research of the thesis. These different stages are further explained in section 1.4. The main research question is as follows:

1. How to design an effective reputation system for an online peer-to-peer platform?

Before this question can be answered, a background research was performed to get familiar with reputation systems. Information about how reputation systems work and how they can be evaluated was gathered by performing a literature study. The following questions were answered in chapter 2:

2. How and why are reputation systems used on online platforms?

3. What makes a good (or bad) reputation system?

In the conclusion of the literature research in section 2.3.6, it was established that a good reputation system can be defined by its credibility and by how effectively it elicits intention to review. Based on this conclusion, the following sub-questions were devised and answered in the discussion found in chapter 7 before the main research question could be answered:

4. What credibility features in reputation systems have the strongest influence on its users?

5. What intention to review features in reputation systems have the strongest influence its users?

6. How can the conceptual framework be refined so it can aid in reputation system design?

1.4 Thesis overview

To determine how a reputation system can be evaluated and what makes a reputation system good, a context analysis was performed. Chapter 2 describes a literature study that discusses related work and provides the metrics that were able to define the performance of a good reputation system. Chapter 3 describes a state-of-the-art study that explores the reputation systems of different popular platforms. During this study, a selection of reputation system features was gathered. The data gathered in both studies was used to design a conceptual framework which is presented in chapter 4. The conceptual framework can aid in evaluating and comparing online reputation systems. This conceptual framework was used as a guide during the conduction of the experiments, and was later improved based on gathered results. All the findings of the context analysis (summarized in the aforementioned conceptual framework which can be found in chapter 4) were used to design the first iteration of a user test. A description of the experiment can be found in chapter 5. The test was exploratory: the goal was to find out how effective the gathered methods to positively influence credibility and intention to review were, and how people perceived these methods. Section 5.3 explains how

6

the results of the first experiment determined the design of the second iteration, as the second test was designed to research the interesting findings of the first test. The second experiment is presented in chapter 6. This experiment focuses on four topics: the distinction in behavior between platform users and non-platform users, the degree of which the credibility of a review is influenced by its textual content, reluctance in online user feedback, and finally, how gamification features are perceived. Chapter 7 evaluates the results that were found during the two experiments and limitations to the study. After new insights were gathered, a revision of the original conceptual framework was made. The updated framework can be used as a guidance tool in future research and in the design of new reputation systems. In chapter 8, the thesis is concluded with the improved framework. The main research question is answered and suggestions for future studies are made, along with the implications of the study.

2 Context Analysis

The goal of the context analysis was to learn how reputation systems work, and to define what makes a reputation system good. This knowledge was used to design a conceptual framework that is able to evaluate the performance of reputation systems of online platforms. To create this framework, an overview of features relating to the performance of existing reputation systems was made. This overview was made by performing a literature and state of the art study regarding reputation systems. It is important to note that this background research was documented in a report called "research topics" before this thesis was written. This background research has been extended and revisited, which resulted in the context analysis that can be found here.

Before the aforementioned conceptual framework could be designed, two research questions had to be answered first. Answers to these questions provided sufficient knowledge that was needed to create a conceptual framework.

1. How and why are reputation systems used on online platforms?

2. What makes a good (or bad) reputation system?

To acquire knowledge relevant to reputation systems, first a literature study was performed to get familiar with the science behind reputation systems. How do they work, how can they be measured in terms of performance, and what has already been researched?

2.1 How and why are reputation systems used on online platforms?

Before looking at what factors influence the performance of a reputation system, it is important to understand what a reputation system is, and what the importance of reputation systems is. In essence, a reputation system can be described as a tool that allows users of a product or service to leave their feedback, often for other users to see. Leaving feedback via reputation systems can be done in many ways, and is often implemented in the form of leaving comments, ratings, or recommendations. Generally, the purpose of this feedback is to generate trust, as online environments lack traditional cues used to evaluate reputation and trustworthiness that we are used to in the physical world (Jøsang et al., 2007).

In literature, a variety of papers can be found that explain in greater detail the importance of reputation systems. In many of these papers, reputation systems are mentioned in the context of online stores and platforms. Online shopping has become a prevalent occuring form of business, which means that purchases are often made by buyers that have not had any form of physical interaction with the product or seller. Users want to base their decision of purchasing a product on more than the description of a product made by the seller, they want to read comments provided by other customers that have bought the same product (Mudambi and Schuff, 2010; Calheiros et al., 2017; Blazevic et al., 2013). Comments of other buyers are perceived as more truthful than the words of the seller. According to a study conducted by Utz et al. (2012) that explored the impact of online reviews on consumer trust, reviews were perceived as the strongest predictor of trustworthiness judgments, even more important than the reputation of the store where products are presented in. When users trust the information that they obtain online, they can be persuaded into purchasing items (Bulmer and DiMaurio, 2010). This is

in line with literature about electronic word-of-mouth (e-WOM) made by Dellarocas (2003). According to Dellarocas, the most important channel for e-WOM dissemination is reputation systems. Reputation systems address the part of human nature that wants to know what our peers are doing. E-WOM plays an important role in consumer decision making, indicating that online consumer communities empower consumers (Costa et al., 2019). Additionally, online reviews are useful to sellers as they provide valuable customer data (e.g. seeing the opinion of a buyer about their products). Therefore, this exchange of feedback in a reputation system context is valuable for online businesses (Hajli, 2018).

Finally, Basili and Rossi (2020) highlight that reputation systems can be used as an effective governance tool, a way for a platform to enforce behaviour within the community which is aligned with the best interest of the platform (e.g. disabling accounts of users that have a reputation that is lower than a certain threshold).

In conclusion: reputation systems are important to generate trust among online buyers and sellers and to stimulate users to engage with each other. Review systems allow e-WOM to occur, enhancing consumer decision making and product awareness. Review systems provide buyers and sellers with information about the product, and can be used as a tool of governance, promoting positive behavior among the users of a platform.

2.2 What makes a good (or bad) reputation system?

This section aims to create a distinction between good and bad practices when designing a reputation system, and to find out what good or bad means in the context of a reputation system. In section 2.3.6 the answer to this question is presented. It is important to note that there are multiple parties that interact with reputation systems on a platform, who all have different interests. Owners of an online platform would consider a reputation system good if it increases their profits (for example by engaging more users to buy products). Sellers on a platform would benefit from a reputation system that enhances their visibility on a platform, to gain more traffic. Buyers on a platform benefit from reputation systems that show products in an honest way, so they can make their choice of buying products in an unbiased way. When looking at what makes a reputation system good, all of these opinions were considered. However, the related literature mainly considered the perspective of the buyers, as they are the largest group of users that interact with reputation systems on a platform.

2.2.1 What makes a system good?

There are many ways to design and implement a reputation system, which makes it difficult to clearly define what actually makes a reputation system good. To get more familiar with what metrics can be used to evaluate the quality of a reputation system, a study of different metrics that are commonly used in the design of reputation systems was performed and is described in section 2.3.

Jøsang et al. writes that the main function of a reputation system is to build trust among users of online communities (Jøsang et al., 2007). Resnick et al. describe three properties that are vital for the performance of a reputation system (Resnick et al., 2000):

1. Entities (agents participating in a reputation system) must have a long lifetime and create accurate expectations of future interactions.

2. Ratings about current interactions must capture and distribute feedback about prior interactions.

3. Ratings about past interactions must guide decisions about current interactions.

The first point mentions the longevity of agents, which means that it should be impossible (or difficult) for an agent to change identity and purposefully remove its connection to past behavior. All of the mentioned properties rely on user feedback, which emphasises the importance of user feedback for a reputation system. It seems logical that without user feedback, a reputation system cannot support an environment of trust. Dingledine et al. (2001) proposed a set of basic criteria for reputation computation engines (the engine that computes the reputation of an individual within a reputation system), to judge the quality and soundness of a reputation system as a whole:

1. **Accuracy for long-term performance:** The system must reflect the confidence of a given score. It must also have the capability to distinguish between a new entity of unknown quality and an entity with poor long-term performance.

2. **Weighting toward current behaviour:** The system must recognise and reflect recent trends in entity performance. For example, an entity that has behaved well for a long time but suddenly goes downhill should be quickly recognised as untrustworthy.

3. **Robustness against attacks:** The system should resist attempts of entities to manipulate reputation scores.

4. **Smoothness:** Adding any single outlier rating should not influence the score significantly.

These criteria are in line with Resnick et al. and focus less on user feedback, but instead introduce additional points like the way a system adapts to a timeframe. This knowledge can help with establishing a general direction for the metric research that is performed in section 2.3.

Looking at the previous mentioned properties and criteria, it appears that user feedback is vital for the performance of a reputation system. This means that it is important for a reputation system to stimulate users to leave reviews, while the purpose of the system itself should be managing the expectations of other users of the platform in an accurate way. The chosen metrics that are used to define how 'good' a system is, should capture these two overarching goals in some way. Another way to look at the metrics evaluation problem is finding metrics and concepts that benefit both the buyer and the seller (and the platform as well). If all involved parties benefit from a system, there will be more reason for platforms to adapt towards this system. To get familiar with more concepts that determine the performance of a reputation system, a variety of problems that occur in current implementations of reputation systems are analyzed below.

2.2.2 Problems in reputation systems

There has been thorough documentation in literature about problems related to reputation systems. Common problems include feedback under-provision, negative-review reluctance, dishonest reports, social influence bias, selection bias and rating inflation (de Langhe et al., 2015; Resnick et al., 2000; Basili and Rossi, 2020).

Feedback under-provision occurs when there are no users that leave a review. While lack of feedback is a problem in itself, it also causes a negative feedback loop: when users see products that have no reviews, they will be less inclined to try out this product and leave a review themselves. This problem occurs when users are not willing to provide feedback, and the option to give feedback is not required. While leaving no feedback (often) creates a more streamlined experience for a buyer, it prevents the forming of an environment of trust and reputation. A phenomenon that causes an under-provision of reviews and relates to the *intention to review* of people is U-shaped feedback: people tend to only leave a review if their experience was either really good or really bad, but the in between area will not always trigger people to leave a review (Meijerink and Schoenmakers, 2019). This phenomenon contradicts slightly with a problem that is called negative-review reluctance: this occurs when people are unsatisfied with a transaction, but due to the public nature of ratings they do not want to leave a negative review for fear of retaliation. If ratings are given anonymously, it can be easier for users to provide honest feedback as there is no fear of their feedback being linked back to them. This does however give less credibility to the given feedback as other users have no way to validate the reviews and see who actually left them.

If people do leave reviews and provide feedback on the platform about other users, more problems arise. This problem is that some people don't leave genuine reviews (dishonest reports). As has been in the Dutch news recently[2], this flaw in current reputation systems on online platforms can be quite troublesome. When looking for services to use, a lot of people trust online reviews (ter Huurne et al., 2017) to guide them, which means that these literally can make or break a business. People can use them to discredit their competitors, or to make themselves seem better then they are. If an attack is launched at a user, many negative ratings can be given to discredit them (this is also known as "Ballot box stuffing"). If people become aware of fake reports this has a negative impact on the trust in a platform, which results in a bigger challenge for the platform to maintain a healthy community.

Other pitfalls to effective reputation systems described by Jøsang et al. include change of identities and discrimination. Again, these ideas tie back to the idea of regulating user actions in order to gain accurate and consistent user feedback. When analyzing different types of reputation systems it is important to look at these specific features in order to determine the effectiveness of each system.

In conclusion: there are multiple ways in which a reputation system can prove to be ineffective. The problems that were discussed in this section have been described in literature, however, the total number of problems in reputation systems is not limited to this selection. However, this section did provide a general view of what can go wrong when designing a reputation system, and it gave information about what to look for in metrics that describe the performance of reputation systems. It is clear that the properties that make a reputation

---

[2]link: nos.nl/op3/artikel/2273850-online-gesjoemel-reviews-eenvoudig-te-koop.html

system good (sufficient user feedback, feedback that elicits trust) are directly connected with the properties that make a system bad (feedback under-provision, dishonest reports). This knowledge was used in the next section, where literature was used to define the metrics that were able to effectively evaluate the performance of reputation systems.

2.3 Literature perspective of metrics

To get familiar with evaluating the performance of reputation systems, a literature study was carried out to get a reference of the metrics and evaluation methods that are used in scientific research to evaluate reputation systems. Metrics used for evaluation will be approached from a literature and a UX design perspective, as both perspectives create interesting insights into what makes a design good.

2.3.1 Evaluating reputation systems

Before research on specific metrics was performed, a more general study was done by researching papers that evaluate reputation systems. Two studies that evaluate reputation systems in the context of online platforms were found, but while one paper explored the technical side of reputation systems (Liu and Munro, 2012), the other paper provided information that was more practical and applicable to this study. Basili and Rossi (2020) performed a multiple case study on the reputation systems of 9 sharing economy platforms (focused on ride sharing). Basili and Rossi mainly explored the effects of the incentive design of reputation systems[3] on the buyers of the platform, and largely disregard the effect that it has on the drivers (Sellers).

The metrics that Basili and Rossi used to evaluate different reputation systems are listed in table 1 below, and can be used to draw inspiration from when thinking about metrics that can be used to define the quality of reputation systems.

The investigated ride sharing platforms seem to have adopted a strategy that does not utilize reputation systems as an incentive device, but rather uses them to punish poor behavior (E.g. if an Uber driver has a low acceptance rate, or a low average rating on their last 500 rides, this can lead to platform exclusion). A proposition that builds on this premise is suggested by Basili and Rossi: they propose that if the rate in which a user is able to restore their reputation on a platform is reduced, their compliance will be increased. Another note that they made is that the reputation of a user directly links to their remuneration, as platform users usually take reputational ratings into account when choosing service providers. By making the drivers aware of this (e.g. increase their fare based on positive ratings), this concept could be presented in a way that creates more intrinsic compliance.

2.3.2 Trust

Trust is a metric that is mentioned often in the context of online reputation systems. While many decisions in our day-to-day life are made based on trust, trust is a metric that is hard to define. To define the base concept of trust, the Oxford dictionary definition of trust is presented

---

[3]Incentive design aims to align the interest of customers to that of a company (e.g. giving the customer a reward if they decide to leave online feedback).

Metric	Clarification
Certification/verification	This adds to the credibility of users: if a platform supports verification, it's easier to observe if a user is 'real', and users that have verified themselves are less likely to behave maliciously since their behaviour and actions on the platform can be linked to them in real life as well.
Nature of ratings	Two crucial distinctions among rating systems explored by the literature concern whether they are one-way or two-way and whether they are open, double blind or anonymous.[4]
Scale of ratings	It is important to consider the scale adopted for ratings and whether ratings are used to provide incentives and/or to exclude users from the platform.
Thresholds for exclusion	How low (and for how long) can a user of a platform be rated before he gets removed from the platform?

Table 1: Overview of metrics used to evaluate ridesharing platforms, Basili and Rossi (2020)

here: *"the belief that somebody / something is good, sincere, honest, etc. and will not try to harm or trick you"*.

Reputation systems are one of the key antecedents of trust on online platforms, although trust is considerably more complex and clearly extends beyond reputation per se (ter Huurne et al., 2017). To get a better idea of what trust entails and how it can be used to measure the performance of a reputation system, different definitions of trust by Jøsang et al. (2007) are explored to find out what definition of trust could potentially fit as a metric.

Jøsang et al. define two common definitions of trust which they call reliability trust and decision trust respectively. These are the most used definitions throughout literature.

1. **Reliability trust:** Trust is the subjective probability by which an individual, A, expects that another individual, B, performs a given action on which A's welfare depends.

2. **Decision trust:**. Trust is the extent to which one party is willing to depend on something or somebody in a given situation with a feeling of relative security, even though negative

---

[4]In this context, double blind means that both the seller and the buyer are asked to review each other. The provided reviews will only be visible after both parties provide a review.

consequences are possible.

When placing these definitions in the context of a reputation system, it appears that Decision trust seems to be a fitting form of trust, as users depend on the truthfulness of existing reviews to make their decision to engage in a transaction with a seller / buyer. Jøsang et al. explain more categories of trust later in their paper. They distinguish between a set of different trust classes according to Grandison and Sloman's classification in "A Survey of Trust in Internet Applications" (Grandison and Sloman, 2000).

1. **Provision trust** describes the relying party's trust in a service or resource provider. It is relevant when the relying party is a user seeking protection from malicious or unreliable service providers. For example when a contract specifies quality requirements for the delivery of services, then this business trust would be provision trust in our terminology.

2. **Access trust** describes trust in higher authorities for the purpose of accessing resources owned by or under the responsibility of the relying party. This relates to the access control paradigm which is a central element in computer security.

3. **Delegation trust** describes trust in an agent (the delegate) that acts and makes decisions on behalf of the relying party. Grandison and Sloman point out that acting on one's behalf can be considered to a special form of service provision.

4. **Identity trust** describes the belief that an agent identity is as claimed. Trust systems that derive identity trust are typically authentication schemes such as X.509 and PGP (Zimmermann, 1995). Identity trust systems have been discussed mostly in the information security community.

5. **Context trust** describes the extent to which the relying party believes that the necessary systems and institutions are in place in order to support the transaction and provide a safety net in case something should go wrong. Factors for this type of trust can for example be critical infrastructures, insurance, legal system, law enforcement and stability of society in general.

All these different forms of trust give some insight into what trust entails, but make it hard to find what forms of trust should be used as a metric to determine the quality of a reputation system, and if one form should be used or multiple. Most of these 'sub-forms' of trust have some part in the overarching concept of trust that is hypothetically usable as a metric. The many descriptions of trust make it hard to quantify trust as a metric and to correctly assume what form of trust is used when looked at it from an evaluation perspective. The multiple interpretations of trust have caused problems in related research before (Thakur, 2018): *"It's difficult when comparing research that has been done about trust, as one stream of research considers trust as antecedent while others consider it as a consequence of customer engagement (Bowden, 2009a; Brodie et al., 2011a, b; Jaakkola and Alexander, 2014a, b; Vivek et al., 2012)."*

In the context of online platforms, the most applicable form of trust seems to be Decision trust as defined by Jøsang et al.. Mutual trust enables collaborative interactions on online

platforms (Hawlitschek et al., 2018; Mazzella et al., 2016). Platforms are heavily reliant on the harnessing of network effects to attract users (both buyers and sellers) to sustain their business model. Because of the low entry barrier for non-professional users compared to a non-P2P service, the quality of service on platforms generally has a wider variance, which can damage trust in a platform. Another big difference compared to a non-P2P service is that the consequences of P2P online interactions can be harsher than in standard online sales. Trusting an unreliable driver on Uber can (under extreme circumstances) cause physical damage to a user, while being scammed by a regular online seller will usually just result in monetary loss (ter Huurne et al., 2017).

Most of these problems can be mitigated by only engaging with trusted users, and effective reputation systems play a key role in this regard (Katz, 2015; Thierer et al., 2015). Mauri et al. (2018) explain that compared to non P2P businesses, branding and brand images are not deemed an effective way to gain trust when users are not professionals. Evidence exists, however, that personal branding plays an important role in boosting popularity in the context of sharing economy applications in the hospitality sector, and standard public regulatory tools meant to ensure consumer protection have not so far been extensively applied to the sharing economy (Schor and Fitzmaurice, 2015).

If reviews are necessary to create more transparency and trust, but a negative reputation can drive away new potential customers, what can online businesses do? Next to promoting and the enabling of personal branding, a way to for businesses to improve is by responding to user reviews. Review Trackers (2018)[5] shows that more than half (53%) of customers expect businesses to respond to their online reviews, and Market Bitz (2018)[6] found that 33% of Yelp reviewers will update their review if a business responds to them within a day. The Review Trackers survey also highlights that 45% of the respondents claimed to be more likely to visit a business if it responds to negative reviews.

These numbers show that customers do value a company or business that listens to their voice and takes action if needed. By responding to reviews, companies can mitigate negative impact of reviews and show their own perspective by contextualizing the experience of a customer. By acknowledging and addressing customer complaints, companies are able to show to new potential customers what the company values are, and that it cares about its customers. The company in this context can be a platform or a seller on a platform, but extends to more online environments.

2.3.3 Credibility

Credibility is a metric that shows many similarities to trust. Fogg and Tseng wrote a paper on 'computer credibility' (Fogg and Tseng, 1999), which offers a solid theory of online credibility and is often cited in literature. According to Fogg and Tseng, the two key dimensions of credibility are trustworthiness and expertise. Fogg defines trustworthiness in the terms well–intentioned and unbiased, and expertise in terms of perceived knowledge, skill, and experience (Fogg, 2003). In addition, Fogg and Tseng stress that credibility is a perceived quality,

---

[5]link: reviewtrackers.com/reports/online-reviews-survey/
[6]link: marketingbitz.com/why-it-is-important-to-respond-to-reviews/

which gives it a subjective nature. This means that we cannot necessarily design credibility itself, but rather design for credibility.
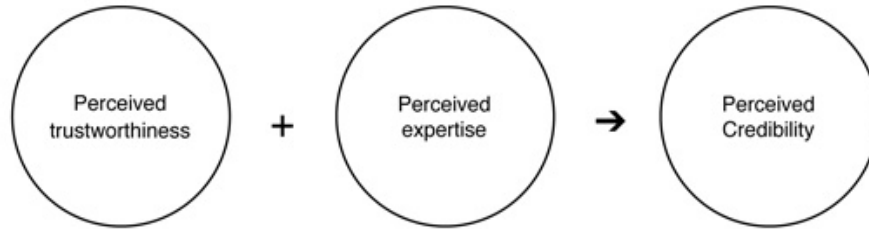

Figure 1: Model of the key dimensions of credibility (Fogg, 2003)

According to the model presented in figure 1, people in an online environment try to estimate credibility by looking for cues related to trustworthiness and expertise. Jessen and Jørgensen (2011) note that this behavior is similar to evaluating the value of any particular source (which also applies to an offline setting); when estimating the value of a source, people look at the trustworthiness of an author and whether the author is an expert on the topic. Evaluating the credibility of an author is easier than evaluating their claims, which makes this an efficient approach.

A problem with this approach is that a lot of information in today's online environment does not come with cues that support perceived trustworthiness or expertise. Online platforms often consist of user generated content, where users can be anonymous (e.g. platforms such as wikis, blogs and forums). If the model of Fogg and Tseng is applied in this context, that would mean that the credibility of these sources would be very low. However, recent studies suggest that information without identifiable authors can still be perceived as credible (Pettingill, 2006; Hargittai et al., 2010), if a collective form of verification exists (Lankes, 2008; O'Byrne, 2009). This means that the feedback of other people is important for determining the credibility of information found online (Weinschenk, 2009; Ljung and Wahlforss, 2008). Jessen and Jørgensen describe these trusted feedback providers as 'trustees' (Pettingill, 2006). Trustees can act as a form of autonomous authority, providing a baseline of trustworthiness to other users. These trustees don't have to be an expert on the subject, but are still important for the overall dynamic of establishing credibility (Wang and Emurian, 2005). This trustee-aspect of reputation systems is mentioned earlier (in section 2.3.1), as Uber uses the rating given by trustees to determine if a driver is allowed to operate on their platform.

Social validation of the quality of information or content is important for assessing credibility. Reputation systems enable users to rate content (using comments, ratings, likes). Jessen and Jørgensen call the collection of multiple streams of trustworthiness cues the 'theory of online credibility: aggregated trustworthiness'. They created an updated model based on the model of the key dimensions of credibility from Fogg and Tseng (see figure 2), adapted to the online environment of today.

The perceived credibility (on the right side of the arrow) is the degree of which people trust the information presented to them. This is influenced by three main factors on the left side:
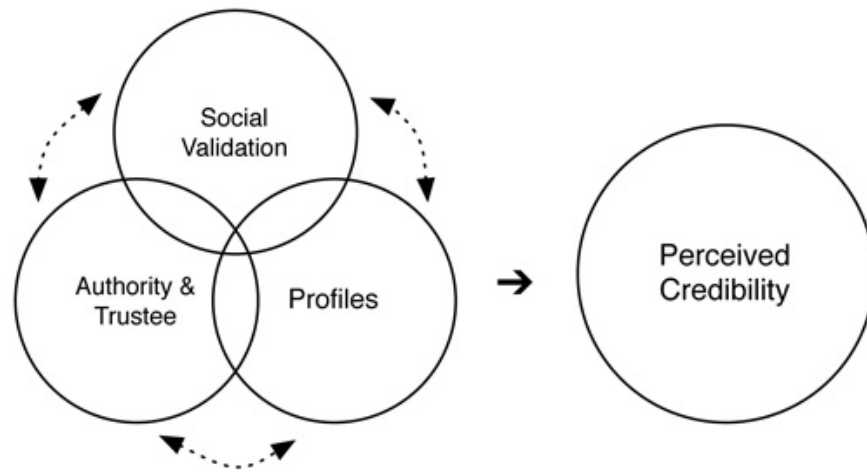
Figure 2: Illustration of aggregated trustworthiness (Jessen and Jørgensen, 2011)

1. **Social validation:** This is the collection of user feedback that is left by people online (e.g. comments, likes, shares, ratings). In general, the more people that acknowledge a piece of information, the more socially validated this piece is (becoming more trustworthy).

2. **Authority & Trustee:** The authority that displays information (e.g. the University of Twente that publishes a paper, or a newspaper that prints articles) influences the degree of trustworthiness of the provided information. Trustees have the same function, but in a lesser degree (e.g. friends on social networks, twitter personas).

3. **Profiles:** These can be a part of social validation. Profiles provide a baseline for identity, and by adding trusted profiles to a user profile (such as a personal facebook, linkedin or twitter), the integrity of a profile can be validated. The identity of a content provider can be important when evaluating the provided content or information of this person.

The dotted arrows and overlapping sections in the figure indicate that all the factors on the left side of the arrow are co-dependent, and imply the navigation between the factors. This dynamic is enhanced by navigation- and search processes (e.g. rankings, user history).

Jessen and Jørgensen based their theory of aggregated trustworthiness on two major studies of young adults' evaluation of Web content and information–seeking routines (Hargittai et al., 2010) and youth's online research practices (Pettingill, 2006). These studies show how youth collects credibility cues from a various amount of sources, not just expert sources. Youth uses feedback and evaluations made by trustees (including members of their online network and people serving a guiding role, like teachers or parents) to successfully complete their information-seeking tasks, without exposure to source credentials or traditional expertise cues. Obtaining online trustworthiness cues is the most robust form of evaluation when author credentials are absent (O'Byrne, 2009).

This shows that to properly evaluate credibility online, understanding social dynamics is more important than the traditional methods of source evaluation. It also explains why platforms like Wikipedia and Twitter can be perceived as credible, despite the lack of known authorities. Evaluating other individual's behaviour online (e.g. comments, likes, shares, ratings), and the way that topics are aggregated (e.g. search rankings, online trends) can provide a baseline of judgment to base individual evaluation on (Weinschenk, 2009). Factoring out one root authority and replacing it with multiple, although less stable sources, spreads out the risk of being mislead online and is easier in an online environment than to establish root authority (Lankes, 2008). This dynamic contrasts with the theory of Fogg and Tseng, as depicted in figure 1.

A vital element in stitching together multiple sources to create a patchwork of credibility is the way that users are able to efficiently navigate online to collect necessary information. Filtering systems like search engines and social recommendations are important in this process (Lankes, 2008; Hargittai et al., 2010). The essence of source verification remains intact, the process is just different.

From a P2P platform perspective, credibility can be applied to all involved parties (platform owners, sellers and buyers). While the personal interests of these parties are different, their need for credibility remains the same. It is important for platform owners that the users of their platform are credible, as a P2P platform relies on interactions between real users. They can use reputation systems to prevent scams and inappropriate use of the platform. For sellers, the credibility of buyers and the platform is important because they rely on these parties to obtain their monetary income. For buyers it is important that both the platform and the sellers are credible, as buyers want to spend their money on a product that is the same as described on a platform.

2.3.4 Intention to review

If an online product has many reviews, the average rating becomes more reliable. People tend to prefer products that have been rated over products that have not yet been rated (Jøsang et al., 2007), because the validation of the ratings adds to the trust that people have in a product. This makes the users' intention to review (which is the likeliness that a user will leave a review after engaging with a service or buying a product) a very critical factor in the performance of a reputation system.

A method of gaining more reviews is by eliciting user feedback. There are three common problems that arise when eliciting feedback (Jøsang et al., 2007):

1. The first of these problems is the willingness of users to provide feedback when the option to do so is not required. If an online community has a large stream of interactions happening, but no feedback is gathered, the environment of trust and reputation cannot be formed.

2. The second of these problems is gaining negative feedback from users. Many factors contribute to users not wanting to give negative feedback, the most prominent being a fear of retaliation. When feedback is not anonymous, many users fear retaliation if negative feedback is given.

3. The final problem related to user feedback is eliciting honest feedback from users. Although there is no concrete method for ensuring the truthfulness of feedback, if a community of honest feedback is established, new users will be more likely to give honest feedback as well.

Incentivized feedback is feedback that is given with an incentive in mind, in case of online reviews often a bonus or reward. Feedback that is incentivized is different from feedback that was not incentivized. (Costa et al., 2019) performed a research to understand the difference between unincentivized and incentivized reviews. Data from the analyzed incentivized reviews was gathered by filtering reviews based on certain keywords that indicate an incentive. According to Costa et al.'s study, reviews that were incentivized tend to be lengthier and charged with more sentiment. An important question that arises when thinking about the difference in characteristics of the different types of reviews is whether the incentivizing of reviews also has influence on the integrity of these reviews. According to Costa et al., the average rating of incentivized reviews is slightly higher than the non-incentivized reviews, as can be seen in the boxplot in figure 3. The blue marking on the boxplot indicates the average scores given. Amazon already acts upon lack of integrity in incentivized reviews. If Amazon detects that a seller presents rewards or benefits to obtain positive (fake) reviews, they sue the company that is paying for the fake reviews and the individuals writing them[7].
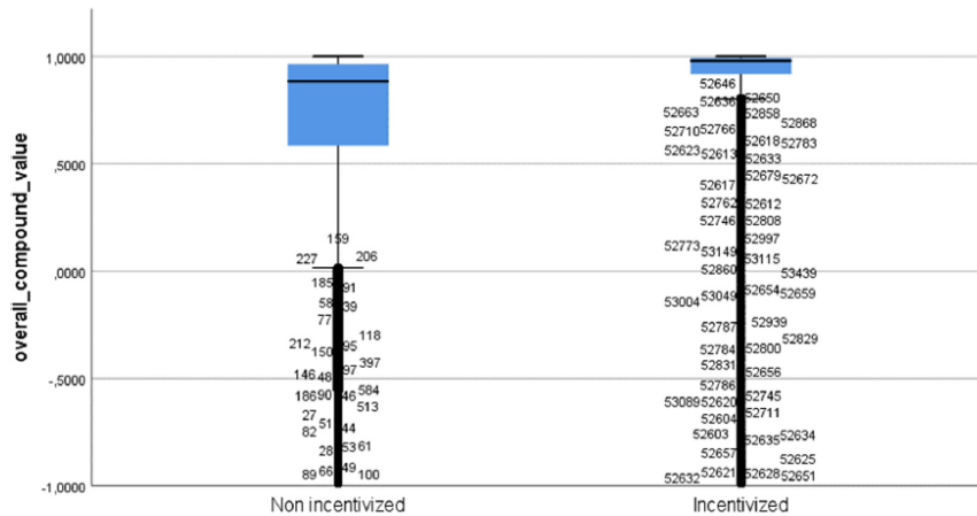
Figure 3: Boxplot of overall score of reviews, non-incentivized vs. incentivized (Costa et al., 2019)

(Thakur, 2018) wrote an article about customer engagement and online reviews in which she designs a conceptual framework that explains the relation between different aspects of an online environment (e.g. an online platform or webshop). Her paper focuses on the way that

---

[7]Amazon.com, 2016. Update on Customer Reviews. Retrieved October 2019, from https://www.amazon.com/p/feature/abpto3jt7fhb5oc.

customers interact with a platform, as well as their likeliness to leave a review. Her conceptual framework can be seen below (in figure 4). Thakur defines customer engagement in this figure as customer engagement with a mobile site/ application of a retailer.
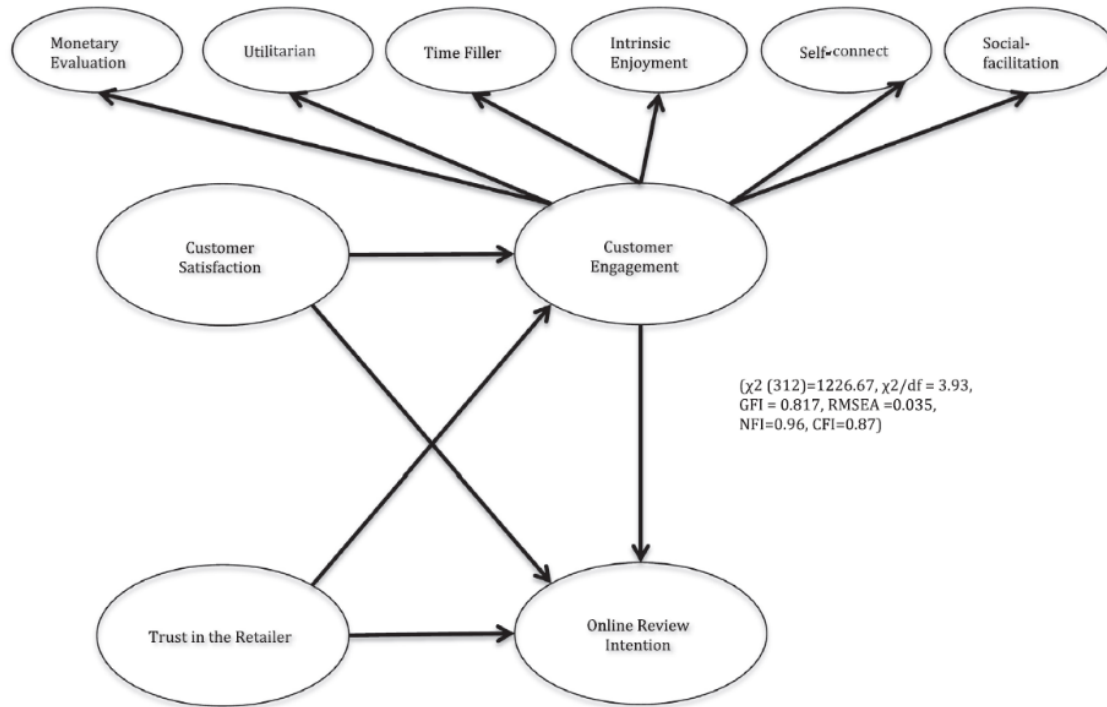

Figure 4: Conceptual framework of customer engagement (Thakur, 2018)

As can be seen in figure 4, all key elements related to online review intention in Thakur's framework directly and indirectly influence the online review intention of users. If *Intention to review* will be the metric that is chosen to evaluate how good a review system is in section 4, the scope of influence should be narrowed to the product itself. Thakur differentiates three main categories that influence online review intention:

1. Trust in the retailer

2. Customer satisfaction

3. Customer engagement

In Thakur's framework, customer engagement is influenced by the customer's interaction with the mobile site/ application of a retailer or platform. The customer satisfaction is determined by the perception of a brand by the customer, and finally the trust in the retailer (as is heavily implied) relies on the users' perception of the retailer. While many previous studies on this subject are conceptual in nature, Thakur provides emperical evidence with a case study that

shows that nurturing the three mentioned aspects will lead to a growth in user reviews on a platform. Thakur highlights that the most effective long term solution to obtain user reviews is to build a relationship of trust with a user. If a user is invested in a platform or product, they are intrinsically motivated to leave feedback. Establishing a positive relationship with customers is an ethical way to incentify users to leave reviews, which is why many platforms aim to establish a positive relationship with their customers.

2.3.5 UX Perspective of metrics

The effectiveness of a reputation system can be described as how much value it brings to the customer and how much value it brings to the business that implements the system. When thinking about what value a reputation system can bring to a user, a UX design model designed by Morville[8] can be used. At the core of UX is ensuring that users find value in what you are providing to them. Morville represents this through his User Experience Honeycomb model, as can be seen below in figure 5.
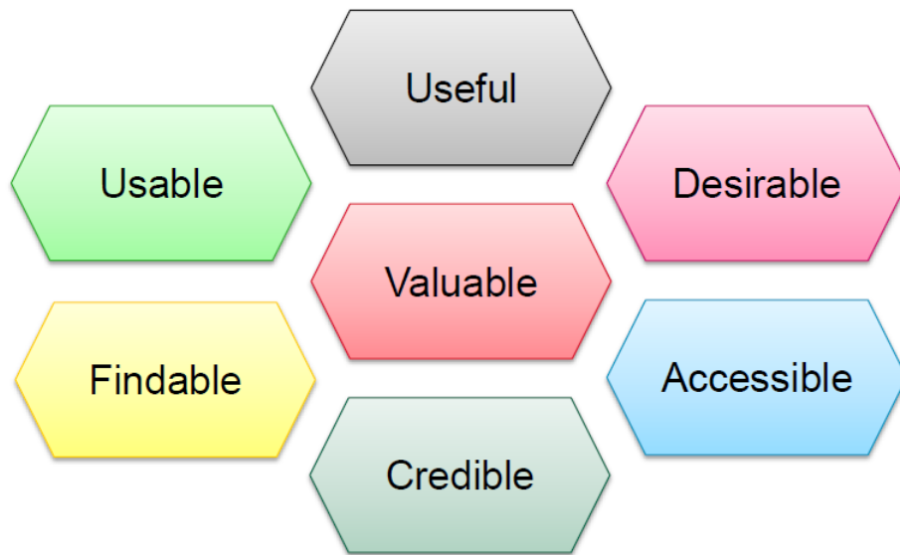

Figure 5: User Experience Honeycomb model (Morville, 2004)

[8]Morville P., 2004. User Experience Design. Retrieved July 2019, from http://www.semanticstudios.com/publications/semantics/000029.php.

The honeycomb model describes different categories of UX requirements to create a valuable tool. Some of these categories focus on making a tool easier or more intuitive to use (which comes down to usability), but if we focus on the categories that care about the actual function and purpose of a tool, four of the mentioned categories stand out.

1. **Useful:** Your content should be original and fulfill a need

2. **Desirable:** Image, identity, brand, and other design elements are used to evoke emotion and appreciation

3. **Findable:** Content needs to be navigable and locatable on site and off site

4. **Credible:** Users must trust and believe what you tell them

It is interesting to see that multiple factors in the Honeycomb model can be related to literature that has been discussed above. The model of Jessen and Jørgensen (figure 2) explains findability as a vital part of enhancing perceived credibility. When Thakur (2018) mentions intention to review in section 2.3.4, it appears that user engagement is influenced by the the same factors that influence desirability in the Honeycomb model (the image, identity and brand).

2.3.6 Conclusion: when is a reputation system good?

Before an answer can be given to the question *"when is a reputation system good?"*, it has to be stressed that there are multiple parties (the platform itself, sellers and buyers) that interact with reputation systems on a platform, who all have different interests. When looking at what makes a reputation system good, all of these interests were considered. However, the related literature mainly considered the buyers, as they are the largest group that interact with reputation systems on a platform. This question is answered with a slight bias towards this group.

When taking a look at all the researched metrics that can be used to evaluate if a reputation system is good, a few things can be concluded:

Trust is vital for the performance of a reputation system. Trust influences whether users will engage with a platform and whether they will rely on the already made reviews within a reputation system. However, trust is difficult to define and to measure accordingly. On top of that, trust can be seen as a metric that is accumulated as a result of an all-round well built system, along with the trustworthiness of the platform itself. As mentioned in section 2.3.2, the multiple definitions of trust have caused problems in related research before. Therefore, it's rather unappealing to use trust as a metric: it's influenced by many factors and hard to measure.

Credibility is a metric similar to trust, but it has a better defined frame and application. There are multiple models with clear antecedents of perceived credibility, Jessen and Jørgensen's model of aggregated trustworthiness found in figure 2 being an example. Similar to decision trust, credibility measures the way that people accept information as truthful and trustworthy. Credibility as a metric covers the expectation management aspect of a reputation system, which was defined as one of the goals of a good reputation system in section 2.2.1. On

top of that, it is a metric that applies well to the context of P2P platforms. All these factors combined make credibility a suitable metric to evaluate reputation systems with.

Intention to review is a metric that can also be considered as vital when measuring the performance of a reputation system. When users have no intention to review there will be no user feedback, and without feedback a reputation system has no value. If a reputation system can maintain credibility while efficiently eliciting user feedback, this would be a highly effective system according to the theory of section 2.2.1. The balance between credibility and intention to review is vital for a good online experience. Frameworks from literature give a clear view of intention to review in an online platform context, as seen in figure 4. Since credibility and intention to review are covering both the expectation management and the stimulation of users to leave reviews, these seem like two fitting metrics that work well together and are able to define the performance of a reputation system. The dynamic of these metrics is interesting, as credibility and intention to review influence each other. A reputation system that is very credible often relies on an in-depth review process, which has a negative impact on intention to review. Finding the correct balance between the two metrics could be vital for the performance of a reputation system.

3 State of the Art

In this chapter an array of online platforms that use reputation systems are analyzed. The strengths, weaknesses and interesting concepts of their respective reputation systems will be elaborated below. During the state of the art research, features of existing reputation systems that influence either the intention to review of users or the credibility of a reputation system will be tracked. All these features will be listed in an overview in section 3.7, which will be used to create a conceptual framework in chapter 4 of this paper. In this chapter mainly peer-to-peer platforms (or P2P platforms) will be discussed. A P2P platform focuses on profitable interactions between groups of users (examples of P2P platforms are Airbnb and Uber). This differs from business-to-consumer platforms (or B2C platforms), where a business often sells their product directly to its customers (examples of B2C platforms are Bol.com or Coolblue).

3.1 Peer-to-peer platforms

A P2P platform usually divides its users into two groups: one type of user is a provider (seller), where the other type of user is a consumer (buyer). There are three types of P2P segments:

1. **P2P-commerce:** online marketplaces where peers can offer or buy goods. An example of such a platform is Ebay or Marktplaats.nl.

2. **P2P-microjobs:** Also known as gig economy platforms, where people offer and request services performed by users. An example of this is Uber.

3. **P2P-sharing:** People that own products or property that is not used can rent these to other users to earn something extra. A famous example of this is Airbnb.

P2P platforms are sustained by interaction between their users: consumers pay for a service or product that is provided by the selling users. As a result, the value of a P2P platform lies in its users and their behavior, which makes the community of a platform very important. A P2P platform has to focus on showing the benefits of mutual interactions very clearly to its users, which can prove itself difficult. In a B2C environment, there is generally one type of interaction (e.g. a user pays a monthly fee, or goes to a store to buy a product), where a P2P platform has to provide multiple ways to interact with the platform or product (buyers and sellers have a very different role on the same platform). In the context of reputation systems, an important difference between P2P and B2C platforms is that P2P platforms have to focus their reputation system more on the interaction between users. With an online community as asset, a great benefit of P2P platforms over B2C companies is that they are very scalable in nature. They do not require direct supervision of a company and are not limited in capacity. Marketing a business to its customers also is different on a P2P platform compared to a B2C company. In section 2.3.2 this difference was already briefly mentioned, as traditional branding techniques for businesses do not work as well on a P2P platform. B2C businesses can market their products more uniformly, as they can assert more control over their product (since it is not as dependent on the varied quality of service within their community).

3.2 General practices online

Some tools that are used in reputation systems are not platform specific. If certain practices prove to be useful or effective, people will be drawn to them, which causes more platforms to incorporate these features. Rating systems that provide a numeric rating about experiences are used on many platforms, a 5-star rating is a common term that people associate with quality. A few examples of rating systems can be viewed in figure 6.
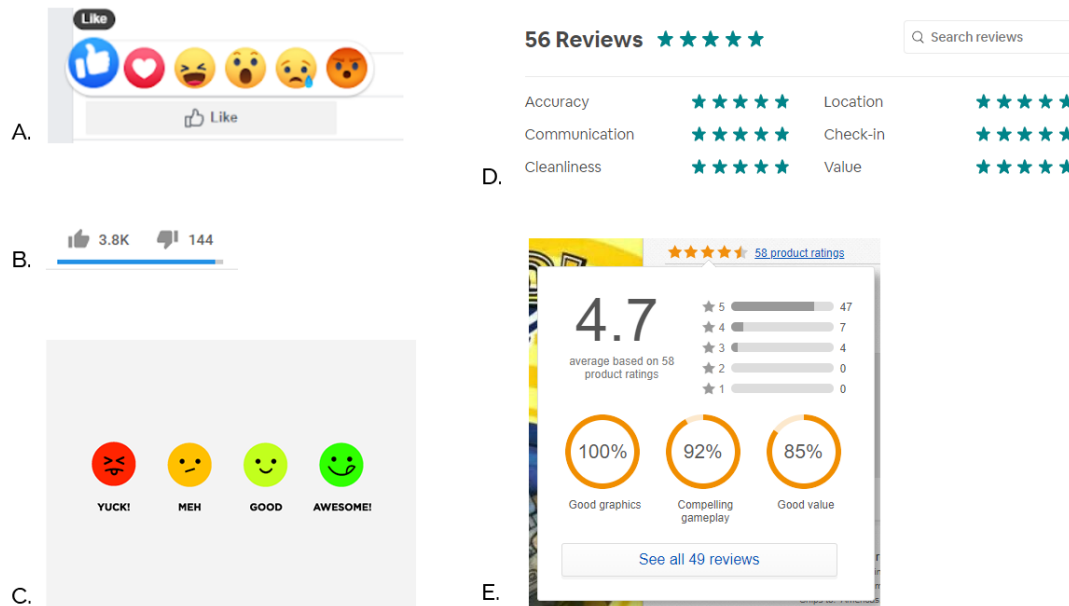

Figure 6: Different examples of rating systems: A. Facebook, B. Youtube, C. rateyour-music.com, D. Airbnb, E. Ebay

Because reputation systems can be vulnerable to attacks that intend to discredit products or users, an often used prevention technique against this is only allowing reviews after the purchase of a product. Another way to prevent bogus reviews is by letting users rate reviews made by other users ("Was this review helpful?").

3.3 AirBnb

Airbnb is a P2P platform where users can sublet their room or apartment to generate an extra income. Airbnb is an interesting case when it comes to review systems, as they are constantly innovating and optimizing their review system. It has become one of their strongest assets, and is supported by an excessive recommender system that uses review data to bring suitable experiences to the right guests. In literature, many papers use Airbnb as an example when talking about P2P platforms. This is why Airbnb is a highlighted example here as well.

Airbnb uses a combination of reputation systems that work together to create a trustworthy experience for their users. There are different types of Airbnb reviews that users can leave on the site or app:

1. **Public reviews:** Up to 500 words that are visible to everyone in the community. Airbnb encourages hosts and guests to leave a public review for each other after a stay; it's not possible to see the review of the other user before you write a review yourself (a two-way double blind review system).

2. **Private feedback:** A message to a host or guest to show appreciation or suggest improvements.

3. **Star ratings:** Ratings for hosts from 1 (worst) to 5 (best) for the overall experience and for specific categories, including: overall experience, cleanliness, accuracy, value, communication, check-in, and location. You need to get 3 star ratings before your overall rating appears on your listing or profile.

4. **Group reviews:** A public review that appears on the profiles of all of the guests on the reservation. There are no group reviews for Airbnb hosts.

5. **Cancellation reviews:** If you cancel a reservation as a host, an automated review will be posted to your profile. These reviews are one of the host cancellation penalties and can't be removed. But you, as a host, can write a public response to clarify why you needed to cancel.

Because the public reviews will be visible to everyone, there is a bias towards positive reviews on Airbnb, as people are afraid that negative reviews will retaliate. This results in a lowering of trust of users in certain areas of the platform.[9] Another problem that Airbnb experiences with their reputation system is that because it is so extensive, it requires a lot of steps to completely fill all the pages (12!) of required feedback. Airbnb's extensive review process turns some customers away from giving their feedback. To give an illustration, the first steps of Airbnb's review system are displayed in figure 7 below.

---

[9]Airbnb community, 2018. Is Airbnb's review system credible enough? Retrieved September 2019, from https://community.withairbnb.com/t5/Hosting/Is-Airbnb-review-system-credible-enough/td-p/768621

Figure 7: The first six steps of the extensive Airbnb review protocol

To encourage good behavior on their platform, Airbnb introduces forms of gamification. An example of this is their superhost system. If a host performs exceptionally well they can become a superhost, which provides a special badge on their profile and gives hosts other benefits, like extra visibility. The current criteria for achieving Superhost status are as follows, as mentioned on their website[10]:

1. Superhosts have a 4.8 or higher average overall rating based on reviews from their Airbnb guests in the past year;

2. Superhosts respond to 90% of new messages within 24 hours;

3. Superhosts have hosted at least 10 stays in the past year or, if they host longer-term reservations, 100 nights over at least 3 stays;

---

[10]Retrieved January 2020, from https://www.airbnb.com/superhost

4. Superhosts cancel less than 1% of the time, not including extenuating circumstances. This means 0 cancellations for hosts with fewer than 100 reservations in a year.

The transparency about these criteria and the prominent display of a Superhost badge on the profile of a host generates trust, and makes a superhost status desirable. This is great for Airbnb, as they enforce positive behaviour in a rewarding way.

3.4 Uber

Uber is a popular P2P ride sharing app where people can order and offer a variety of rides. Uber drivers are generally regular people with extra time and a car that want to generate extra income, but with Uber's rise in popularity, professional taxi companies have started to offer their services on the platform as well.

Uber created commotion as a P2P platform due to their competition with the existing taxi industry, since they did not have to adhere to certain laws (as there were previously no laws for online platforms that did exist for taxis), which allowed their service to be cheaper and less professional than was acceptable in the eyes of existing taxi companies. This created commotion in the taxi industry and the urgency for new laws to regulate ride sharing services that do not enforce a taxi license.[11]

Another controversial aspect about Uber is their lack in transparency about their ride finding algorithm, as well as their nontransparent price calculation for drivers. The lack in transparency causes both benefits and negative consequences. A positive side is that certain drivers try to optimize their platform behavior by trying to drive at different times, in different districts and with different kind of cars (just to find out what will yield them the best results while doing their work). This is a benefit for Uber, as it results in extra supply for the platform. A negative effect of the lack of transparency is that not understanding the platform refrains people from trying their service (either as a rider or a driver). If changes in Ubers' algorithms are made this will not immediately be clear to users, which can cause unwanted situations (e.g. drivers that suddenly earn less or do not get offered any pick-ups). A phenomenon that was talked about in section 2.2.2 is negative-review reluctance. This is a problem on Uber: it is interesting to see that due to the naturally high ratings, a user rating of 4.23/5 stars can already be seen as exceptionally low[12]. Uber drivers are reluctant to pick someone up with ratings below 4.5, while in a traditional sense of rating this is an exceptional score.

3.5 Wikipedia

Wikipedia is a platform that serves as an online Encyclopedia. Articles on Wikipedia are written and rewritten by multiple contributors, of which very few are known authorities on the matter. As talked about in section 2.3.2, traditional methods of source validation are not viable in a setting like Wikipedia, which calls for a different method to validate credibility. A method of

---

[11]Openbaar Ministerie, 2019. Uber pays more than 2.3 million for violating Taxi Act. Retrieved January 2020, from https://www.om.nl/@105314/uber-pays-more-than/

[12]Waters C., 2018. I like to think I'm a nice person but my Uber rating tells a different story. Retrieved January 2020, from https://www.smh.com.au/business/small-business/i-like-to-think-i-m-a-nice-person-but-my-uber-rating-tells-a-different-story-20180806-p4zvqc.html

validation is based on the concept that we do not trust the individual user, but we do trust a lot of them (Lankes,2008). The combined number of edits and re-edits from the large amount of anonymous contributors seem to justify the lack of their authority and identity. A reason for this acceptance could be that it would be extremely cumbersome to coordinate or influence the actions of a large number of people as an individual, especially when all performed actions are transparent. It is unlikely that an author bribed all other editors to make a contribution, the chance is higher that the editors independently agreed upon the contribution of the author.

This dynamic of many anynomous users that contribute to Wikipedia lead to the development of Wikitrust[13]. Critics of Wikipedia argued that there is no easy way to see which articles or what part of an article is credible, and which part is not (Stvilia et al., 2005). To counter this critisicm, researchers developed WikiTrust, a tool that is able to visually show the credibility of an article based on the reliability of the author and the amount of time that edits have existed on the page. The number of edits made are tracked and used to give a credibility value to an author. The more unedited pieces of text the author has written, the higher their reputation. Text written by unverified or questionable sources is coded in orange, which will eventually turn to lighter shades of orange after more edits are made while the text persists on the page (see figure 8). A negative side effect that is caused by this ranking method is that authors that contribute to controversial pages on Wikipedia are more likely to be ranked as 'untrustworthy'.

Open platforms like Wikipedia and the addition of algorithms like WikiTrust are good examples of aggregated trustworthiness in an online (and mostly anonymous) environment. WikiTrust is one example of an online service that allows people to estimate the credibility of online content.

3.6   Interesting trends in reputation systems

Some companies try to be creative with their rating systems. Medium's review system conveys the real life "review" of an applause to an online article. A user is able to give up to 50 claps to an article to express his gratitude.

A trend that has risen over the last few years with the growing importance of online reputation is the **decentralization of reputation systems**. There are several companies (e.g. Connect.Me, TrustCloud en WhyTrusted) that have acted upon the demand for an online guarantee of trust. These companies focus on collecting data and tracking online behavior to make a profile of someone. These profiles can be used to judge someone's integrity, or to verify if someone is a real person. TrustCloud generates a trust score based on made transactions and reviews that can be shown to peers to prove someones integrity. The Dutch government created an online tool to verify yourself with an online identification (DigID) on webservices that require sensitive personal data. Setting up a personal DigID requires valid legal documents, making this a trustworthy tool for many users. The extra steps of verification supported by real life identification allows a user to feel safe.

Anonther interesting trend which is already mentioned in section 3.3 is the **gamification** of a platform. If buyers get benefits by providing ratings and content, this can create an urge for users to leave reviews. Similarly, if sellers provide an exceptional service, there can be benefits for them. The example mentioned in Airbnb's context was the Superhost system, where

---

[13]Wikitrust - http://wikitrust.soe.ucsc.edu/

Monarch consults the will of the people, represented by parliamentary leaders, in determining who should hold the office. As always, the person who has the broadest support from the members of parliament is chosen by the Monarch and confirmed by a vote of confidence by the Folketing. However, before the parliamentary confirmation, the Prime Minister-elect together with the leaders of his coalition partners selects the other Ministers which make up the Governments and acts as political heads of the various government departments. Cabinet members are occasionally recruited from outside the Folketing.

Since 27 November 2001, the economist Anders Fjogh Rasmussen has been Prime Minister to Denmark.

As known in other parliamentary systems of government, the executive, i.e. the Government, is answerable to the Folketing. Under the Danish constitution, no government may exist with a majority against it, as opposed to the more common rule of government needing a majority for it. It is because of this rule, Denmark often sees minority governments.

Under Prime Minister Rasmussen, Denmark has supported many of the United States foreign policies

Figure 8: The Wikipedia page "Politics of Denmark" as captured by Wikitrust. Unverified and questionable content is highlighted in orange, as can be seen in "Anders Fjogh Rasmussen".

sellers are able to get a special badge along with accompanying privileges when they behave exceptionally well. Methods like these allow platforms to create a platform culture, and by promoting certain behavior they can influence their community without explicitly telling them how to behave. Platforms like YouTube implement gamification in a different way: a Youtube Channel is generally ranked by how many people are subscribed to their content. The big red button on the right top of a Youtube Channel page shows the current number of subscribers of a channel. A bigger number means a bigger audience, which generally indicates a certain quality of content. Many channels have adapted certain subscriber milestones where a special celebration video is made after a subscriber treshold is reached. By making the subscriber count a prominent feature on a Youtube channel, Youtube created a platform culture where content providers care about how many people are following them, which creates traction towards the Youtube platform created by their content creators. Gamification can be a powerful tool to create an intrinsic motivation within users to adhere to the platform culture, which can be used to create more credibility or intention to review (e.g. on Youtube: "Leave a thumbs up if you like this video!").

30

Figure 9: Users of Medium can show their gratitude by giving an article 'claps', Medium's rating system that resembles an applause.

3.7 Conclusion

There are many ways that online platforms are implementing reputation systems. The features used in reputation systems that were found during the literature and state of the art research were combined into an overview as seen in table 2 below, where it is clear to see which features add to what aspects of a reputation system. The found features will be used as a foundation for the conceptual framework that is created in the next chapter.

Platform	Category	Feature	Description	Result
Non-specific	Intention to review	Feedback is part of the user-flow	A platform incorporates their review process in the ordering / usage of their products	Leaving feedback is easy for users if it is part of a process that they are already engaged in
		Prompt that asks for feedback	Send user message after a while: "do you want to leave your opinion on this product?"	Users will be reminded to leave feedback, at the cost of annoying them
		Share a link on social media	Users can share online content on their social media accounts	Word of mouth recommendations are powerful tools to create trust, low effort required from users
		Rating system: 1-5 stars	Users are able to provide their feedback on a scale of 1 - 5	Users can easily give feedback and see feedback of others
		Rating system: 1-10 rating	Users are able to provide their feedback on a scale of 1 - 10	Users can give feedback and see feedback of others
	Credibility / ITR	Rating system: thumbs up / thumbs down	Users are able to provide their feedback on a binary scale (bad / good)	The easiest way to rate a product, there is no room for details in rating
		Rating system: using moods	Users can pick a mood out of a few options (e.g. happy, sad, scared)	The rating becomes more emotion based, it is a novel way of reviewing
		Review transparency	Users can see when reviews were made, if ratings have changed, how is a rating generated	People are able to make a better estimation of the value of reviews if they know how a rating came into existence
		Social media login option	Users can create an account using their existing social media accounts	Low effort for users, creates credibility due to profile link, often offers easy sharing options on linked social media account
		Rating system: comment as review	Users can type about their experience and leave this as a review or comment	More qualitative than a numeric rating, but less quantifiable than numeric values
		Rating system: display amount of ratings given	Users can see how many people have already given a rating before them	More ratings correlates with a higher trustworthiness in a rating
	Credibility	Overview of given ratings	There is a display of the different ratings that are given that have lead to a displayed average	People can make a distinction between the extreme reviews (5- and 1-star ratings) to estimate the value of a product
		Sort existing reviews	Users can sort reviews (e.g. by rating / helpfulness)	People can find reviews that are interesting to them more easily
		Rate existing reviews	Users can give their feedback on existing feedback (e.g. "Was this review helpful?")	The more positive feedback a review has, the more credible it becomes
		Leave review only after purchase	Users can not rate a product without buying it first	Prevents bogus reviews of people that have not used the product
		Link to social media account	Users can provide a link to their social media profiles	It is easier to trust feedback if there is a person that can be held accountable for their words connected to said feedback
		User reviews: Anonymous	A platform allows users to give anonymous feedback	Low treshold for users to provide feedback, but feedback is less trustworthy
		User reviews: Real name	A platform encourages users to use their real name when leaving feedback	Higher treshold for users to provide feedback, but feedback is trustworthy
		Decentralized reputation systems	A platform uses an external party (e.g. DigID) to verify user authentication	Extra steps are needed to authenticate users, but users are more likely to be real people
		Social presence	A platform emphasizes the presence of other online users	If there are other users online that use the same platform, the platform becomes more credible
		Chat functionality	Users are able to get in contact via instant messaging	Chatting with other users allows users to verify comments or ask questions
		Pictures added to reviews	users can add pictures to their reviews	If a review has a picture added to it, it enhances the credibility of the review
		Identification Verification	Users can upload their identification to get a 'verified' badge on their profile	Proves that a user is real with provided identification
P2P specific	Intention to review	Double blind reviews	Users review each other, they cannot see the review of the other user before they write a review themselves	Encourages users to leave a review, prevents bias due to a negative review from another user
	Credibility / ITR	Steps per review process	The review process is divided in multiple stages, elaborating on the details of an experience or product	Feedback is more detailed and specific, but it costs users more effort
	Credibility	Categories of ratings	Users can rate all aspects of their experience (e.g. cleanliness, host behavior), not just the overall experience	More in depth information about a product or experience, requires more effort for users
		Superhosts (Airbnb)	Hosts that perform exceptionally well get a special badge on their profile and appear higher in search rankings	Users are more inclined to go with superhosts if they have multiple options
		Reputation threshold for platform use	Users get removed from a platform if their rating falls below a threshold	Users tend to behave properly if they know their actions have consequences
		Combining reputation systems	Multiple reputation systems are combined (e.g. rating for an Airbnb stay and a rating for the host)	Users can give layered and specific feedback, but it requires more steps

Table 2: Overview of acquired credibility and intention to review (ITR) features

4 Conceptual Framework

This chapter finalizes the context analysis by presenting a conceptual framework. The designed conceptual framework categorizes the performance of a reputation system into either intention to review or credibility. The frameworks seen in figure 2 (Fogg, 2003) and figure 4 (Thakur, 2018) served as a foundation for the conceptual framework below in figure 10. Features found during the state of the art and literature study (for more details about these features, look at table 2) are directed towards different mediating factors in the framework. Some features influence multiple mediating factors, but due to readability only one mediating factor is selected per feature. Finally, these factors then direct towards intention to review or credibility. Trust is represented as a factor that influences both credibility and intention to review: both metrics reside within a body of trust. While trust is hard to define and dependent on the environment a reputation system exists within, it has an impact on the performance of a reputation system, which is why it is shown in the framework.
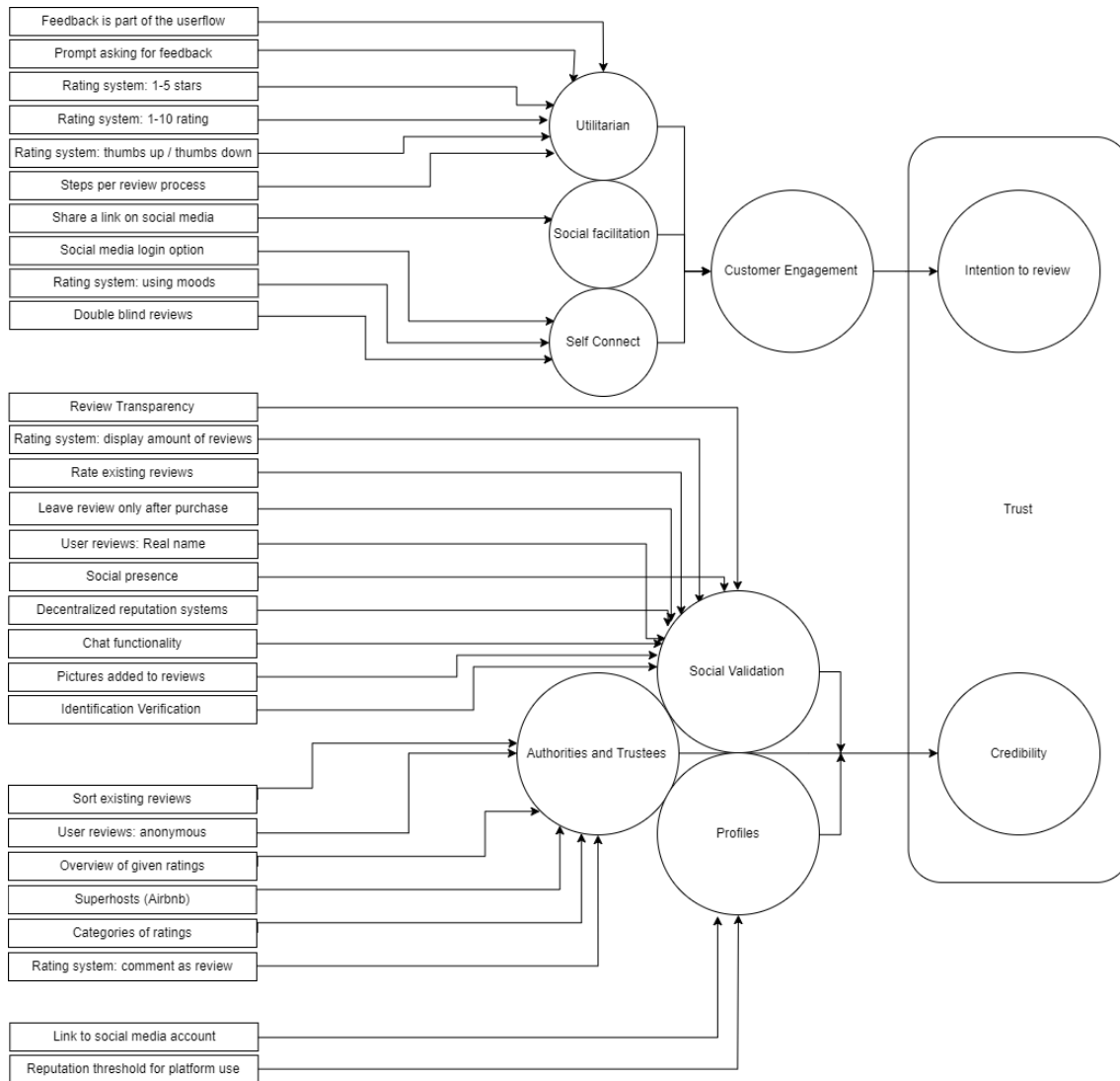
Figure 10: Initial design of conceptual framework

The conceptual framework was designed based on the performed context analysis and can be used to identify reputation system features. The features listed in the framework are used to elicit credibility and intention to review by reputation systems. However, the impact of individual features on the performance of a reputation system is not apparent by looking at the framework. Some of the listed features are used on platforms but are not proven to be effective. An experiment was conducted as described in chapter 5 to verify the performance of the features shown in the framework. Different methods of testing the impact of reputation system features were used, to eventually improve the framework and design a final version. This improved version can be found in section 8.2.

5 Experiment 1 - Framework driven

The information gathered during the context analysis phase was used to design a conceptual framework (chapter 4). This framework was used to design the first experiment. Results from the first experiment were used to design the second iteration of this experiment, allowing the exploration of interesting findings. The end product of these experiments was an updated version of the conceptual framework, that can be used to develop a reputation system based on the obtained design principles. This framework can be found in section 8.2.

The experiments that were performed were designed to be exploratory. The reason for this is that a research concerning the performance of a reputation system based on credibility and intention to review has not yet been performed. The participants of the experiments had to interact with reputation systems from different perspectives: the perspective of a visitor / buyer on a platform and the perspective of a platform user that bought a product and is asked to write a review. Reputation system features were tested by creating multiple custom reputation systems that presented specific system features to participants. After researching the impact of the features in their online context, the features were also tested without their context, in both quantitative and qualitative ways.

The aim of the first experiment was to obtain a better insight in how people online experience specific credibility or intention to review triggers. The set-up of the experiment was dictated by the structure of the conceptual framework. The reputation system features that can be found in the framework served as a source for features that were tested. During the experiment, participants were presented with different reputation system features (both in a platform and non-platform context), and were asked to rate these features. The presented features comprised a multitude of credibility and intention to review triggers. In the first part of the experiment, participants had to pick a coffee machine from a set of coffee machines, all with different reviews. The presented reviews had different credibility features, to learn if credibility features were able to influence the behavior of participants. After this implicit way of testing features, they were tested explicitly. Participants had to rate the importance of various credibility features for the credibility of a reputation system, and they had to explain how different intention to review features impacted their intention to review. The experiment was divided in four sections: credibility features (implicit), credibility features (explicit), intention to review features, and a reflection on intention to review. The experiment was accessible via a web browser, to allow users to remotely perform the test.

5.1 Method

The method section is divided in a few key parts: Participants, Materials, Design and Procedure. After these parts are discussed, the different segments of the experiment are explained in greater detail. The segments are explained in a chronological order.

5.1.1 Participants

In total, 29 people participated. Of these people, there were 16 males and 13 females. For this experiment no specific demographic was selected, with the exception of all participants

being mature and competent. There were participants from different age groups, professions and countries. There was one common factor that connected these people: they use social media, and therefore have some experience with reputation systems. The participants were contacted via Whatsapp, Facebook and by recruiting volunteers in public university spaces. The incentive of participants to partake in the experiment was based on goodwill. The non-specific demographic was chosen to create an honest selection of people online, with no biases involved in the results. All participants received the same test, there were no alternative versions or variables. The wide selection of participants allows to show insights and differences between demographics. The age distribution of participants looks as follows:
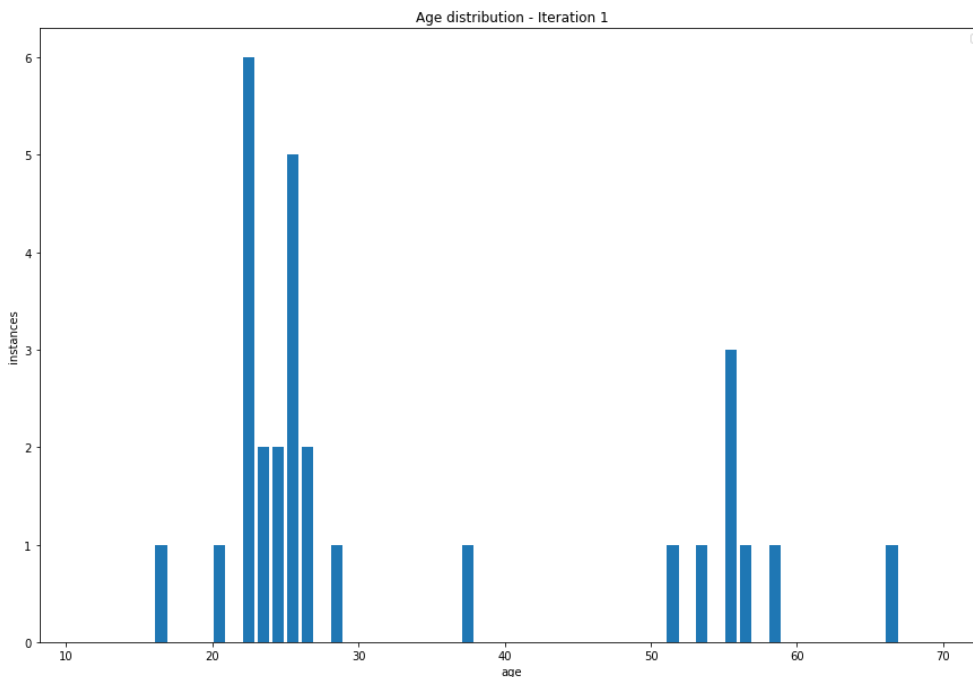
Figure 11: Age distribution plot of the participants of experiment 1

Interesting to note in this distribution is that there are two clusters of participants: There is a group of participants in their mid-twenties and a group of people in their mid-fifties. The difference in behaviour between these age groups could explain how growing up with social media and online technologies can influence the experience of a reputation system.

5.1.2 Materials

One of the requirements of the experiment was for it to be easily distributed online. The benefit of this approach is the bigger reach of people that could be contacted via social media and then participate. This lead to the experiment being designed for the web, being coded in web languages such as html, css and javascript. The experiment was designed for a desktop

computer, but will be functional on mobile phones to reach the highest amount of participants. Every device that can run a web browser was able to run the experiment.

The lay-out and visual design was based on the Listeo theme[14], which was also used as the foundation of the Picobelly website. It is a neutral and clear visual theme that resembles an online environment where reputation systems are used.

For the first part of the experiment a selection of pictures of similarly looking coffee machines was gathered from Tweakers.net to be presented to users as stimuli. These were accompanied by a selection of coffee machine text reviews that were found on Amazon. The text reviews were adjusted later to make them more similar in tone and length. The combination of these stimuli were presented as if they were online product reviews from actual users. In section 5.1.5 this is explained in more detail.

5.1.3 Study Design

As mentioned in the beginning of chapter 5, the experiment was designed to be of an exploratory nature. The goal was to establish the influence of different intention to review and credibility features on users. This means that there were two main dependent variables: intention to review and credibility.

The independent variables were the reputation system features that were presented to users during the user test. These were represented by both credibility - and intention to review triggers. First, credibility features were tested in an implicit way. After that, they were tested in an explicit way. This was done to find out if people react to reputation system features in the same way as they claim to value them when the features are not emphasized. The features that were used were derived from the conceptual framework of chapter 4. The conceptual framework contains a large array of different features that (hypothetically) will have an effect on the user experience (and effectiveness) of a reputation system. To find out what the impact of these individual features is, they have to be measured individually. Testing the individual effectiveness of 33 features requires a significant number of participants and a complex user test (as these features all need a fitting context to be effective). Due to the scope of the thesis, a selection of features was made. This selection encompasses a broad scope of the listed features, while not all features have to be tested. To make a fair selection, the available features were divided into groups that cover a specific part of the reputation system. An example of a group of features is rating systems. Different ways of giving a rating can be tested, but not on the same review. Making a pre-selection of feature types that could be tested brought down the scope of the experiment and yielded more effective data with less results.

Other independent variables were age and gender. These variables were asked from the user before participation, but were not the focus of the research. However, if any interesting findings might happen between age groups and genders, these could be detected.

The controlled variables were the browser environment, the order of the questions during the experiment and the information that was presented during the test.

The experiment used a within-groups design. All participants received the same triggers, and were able to express their reactions that acted upon these triggers. The reason for using

---

[14]Listeo: https://themeforest.net/item/listeo-directory-listings-wordpress-theme/23239259

within-groups design was that by exposing all participants to all different triggers, individual differences would not distort the results. If between-subjects design was used, there would be a possibility that there were differences between the groups that could impact the results of the experiment. On top of that, having within-groups design proves to be effective with small groups of participants.

5.1.4 Procedures

In this section, the procedures performed during the experiment are explained.

The planning of the research: To give a realistic reputation system experience, mild deception was used during the experiment. If specific reviews or reputation system features are presented to users, they need to experience them as if they were real to get accurate results. To legally perform a mild deception in the experiment, the EEMCS Ethics Committee (the ethical committee of the faculty) was asked for permission. This permission was granted.

Set-up of experiment: This experiment was set up in multiple sections. Every section tested either credibility or intention to review triggers. During the experiment, users get presented with different situations that they need to evaluate, and then get corresponding questions they need to answer. The overall experiment is designed like a survey, to make gathering the user data efficient and quantifiable.

Preliminaries of the experiment: Before the experiment started, users were presented with information about the study. The introduction did not mention credibility or intention to review yet, to not prime the user. The study information can be read here:

> Hi, I'm Floris. Thank you for showing interest in my research! The goal of this research is to gain more insight into how people behave in an online shopping environment, with a focus on how people perceive and write online reviews. With the obtained data, more insightful review systems can be designed in the future. The obtained data will be anonymous, the only personal data needed is your age (you must be 16+ years old to participate) and your gender.
> If at any point during the user test you wish to stop participating, you can just close the webpage. You don't have to give a reason for this.
> There are no risks in participating, this research project has been reviewed and approved by the EEMCS Ethics Committee. If any questions arise, feel free to contact me (Floris Veldhuizen, at f.l.m.veldhuizen@student.utwente.nl) or my supervisor (Dennis Reidsma, at d.reidsma@utwente.nl). Finally, you could also reach the Ethics Committee at ethics-comm-ewi@utwente.nl.

After reading the study information, users were prompted to give their informed consent before the test starts. This was necessary because informed participation is an ethical and

legal requirement for research involving human participants. When consent was given and the user is older than 16 years old, the participant could proceed and start with the experiment.

In the next sections, the different parts of the experiment are explained in more detail.

5.1.5 Implicit credibility features

This section is the start of the experiment. It is designed to present credibility features in a natural way, without users being primed towards credibility. Different reviews that all contained unique credibility features were shown to users. The reviews that were shown to the user were presented as if they were real, which is why a small deception was needed. The used features were selected based on groups of features from the conceptual framework of section 4.

The provided context of this section incentivized users to pick a coffee machine out of a selection of six different coffee machines. The presented products with respective reviews were very similar in nature: the designs of the different machines were selected to look similar, the price of the machines was equal and the content of the reviews was all notably positive. The presented reviews were designed to be similar to put the emphasis on the credibility feature. This ensured that the only apparent independent variable was the presented credibility feature. The provided prompt is presented below:

> You have the money to buy a new coffee machine. You can pick any machine out of the six machines presented below. Which one would you pick? A selection of machines and reviews gathered from different online platforms is presented below.

The hypothesis for this section of the experiment was that users would be likely to pick a favorite machine based on what they feel is the most credible review (with the accompanied credibility feature). To not overwhelm users with reviews, a selection of six features (one for every review) was made. All the presented reviews can be found in Appendix A. An example of a review is presented below:
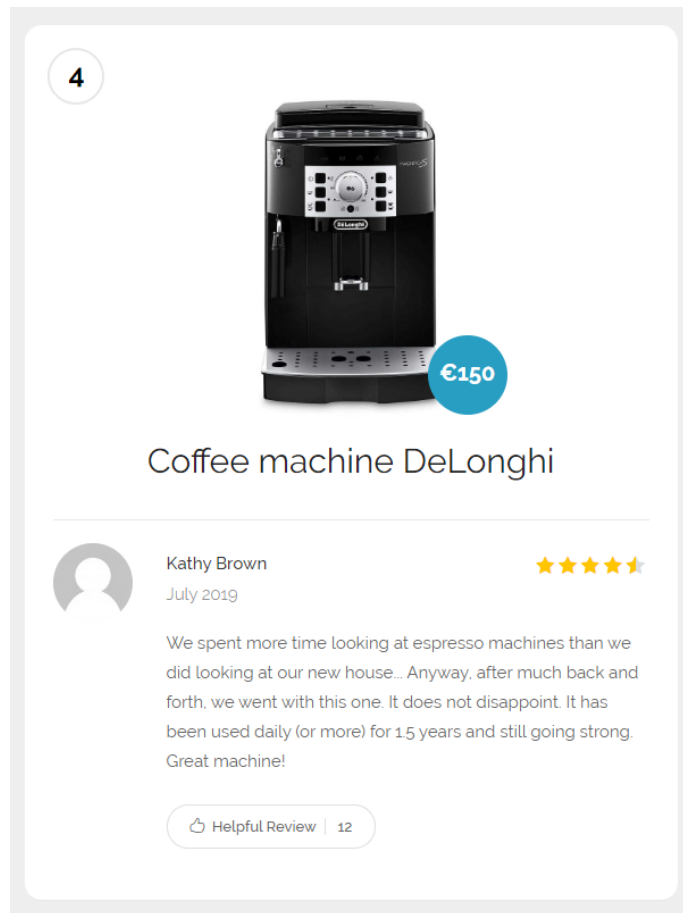
Figure 12: Example review in iteration 1 of the experiment.

In figure 12 one of the credibility triggers is shown: user feedback. This is shown with the "Helpful Review"-button, which allows users to see feedback from others to validate a review. This was one of the tested credibility features. All credibility features that were tested are listed below:

1. Having a link to personal social media

2. Being a verified user (Verified as being authentic by a platform)

3. Being an endorsed user (E.g. an Airbnb 'superhost')

4. User feedback

5. Including a picture in the review

6. Having a personal avatar

The initial plan was to experiment with more credibility features here. Examples of these features are the effect of using a nickname instead of a real name, or seeing the effect of different kinds of rating systems. However, these features caused problems. A link to someone's social media would not be realistic if there was no full name or avatar present to go with the link. If a realistic social media link was added to a review, it would have to add multiple credibility features at once (e.g. a full name, link to a social media platform, possibly an avatar). Therefore, a choice was made to keep a standard selection of features among all reviews. They comprise a rating-system (a 5-star system), the full name of the reviewer and an anonymous avatar. This means that the addition of a link to social media will only introduce a single credibility feature to a review, which gives its impact an honest balance compared to other credibility features that are tested.

Finally, after a participant made a choice out of the reviews ("Which coffee machine would you buy?"), they were asked why they made this specific decision. This might contribute towards clarifying whether it the trigger was the decisive factor for picking a review.

5.1.6 Explicit credibility features

In this section of the experiment, credibility features were tested explicitly. This means that the features were not presented in the context of a review; instead, participants were asked about the features directly via questions. Because of the explicit nature of this section, more credibility features could be tested compared to the previous section. Another benefit of the explicit way of presenting these features was that the results of this section could be compared to the results of the first section (with implicit credibility features). The results of this section were able to show if people react to credibility features in the same way as they claimed to value them in the first section. To aid in the rating of credibility triggers, the definition of credibility was explained to participants with the paragraph that is shown below:

> Credibility is an important metric that depicts how truthful presented information is. Researchers define credibility as "the degree to which people trust the information presented to them". Online platforms and webshops try to present their information as credible as possible by adding extra information and features to reviews. Below I will ask you to rate different features in terms of credibility in an online context.

Users were able to rate how important they felt that certain credibility features are on a Likert scale. The questions that were asked are listed below:

1. How important do you think that seeing someones' full (real) name is for their credibility?

2. How important do you think that seeing someones' nickname is for their credibility?

3. How important do you think that a link to someones' social media is for their credibility?

4. How important do you think that a real picture of someone (as avatar) is for their credibility?

5. How important do you think that being a verified user on a platform is for someones' credibility? (A verified user is verified as being authentic by a platform.)

6. How important do you think that being an endorsed user *(E.g. Airbnb's 'superhosts')* on a platform is for their credibility?

7. How important do you think that user feedback on reviews *(E.g. 5 people found this review helpful)* is for the credibility of a review?

8. How important do you think that a response on a review from the seller is for the credibility of a review?

9. How important do you think that adding a picture to a review *(E.g. "this is a picture of how my box arrived")* is for the credibility of a review?

A factor that was taken into account was that some people value certain credibility features based on how comfortable they are with sharing these features themselves. For example, if someone prefers not to share their own full name, they will not feel that the full name of other people is as important for their credibility. To verify if this was the case, these last questions were asked:

1. How important do you think your own privacy is online?

2. How did your concern for your own privacy influence the answers you gave above? Please elaborate.

5.1.7   Intention to review features

In this section of the experiment, intention to review features were tested. The most realistic intention to review test scenario would be to test features after a participant buys a real product. However, within the scope of the thesis it was not possible to accurately recreate the act of buying a real product and then testing it within a web environment. Instead, participants were provided with an explanation of the intention to review metric, along with an online platform context. The shown text can be found below:

> Intention to review is a metric that defines how much a user wants to leave a review after using an online service. Imagine this situation: you rented a room on Airbnb. You spent your entire weekend in another city and have enjoyed your stay. Once you return, you spend a night at home and while sitting on your couch, you check your phone.

Airbnb was used as an example of a platform because many people are familiar with the service. After this introduction, participants were presented with multiple questions. Every question was followed up by these multiple choice options: "No", "That depends" and "Yes". After answering the multiple choice question, participants were asked to elaborate on their choice. The questions can be found below:

1. A notification from Airbnb opens up: "Help us out, please leave a review!". Would this trigger you to leave a review?

2. Forget about the notification. Instead of a notification, you know that Airbnb will offer you a discount on your next booking if you leave a review after your stay. Would this trigger you to leave a review?

3. Instead of the previous options, Airbnb offers you platform credits that show on your profile after leaving a review. These credits could earn you different badges and platform privileges *(E.g. more accommodation options in the future)*. Would these credits trigger you to leave a review?

4. Instead of the previous options, Airbnb sends you a notification that lets you know that your host has left a review for you. You can only see what your host wrote about you if you leave a review too! Would this trigger you to leave a review?

These questions were designed to ask about different intention to review triggers: either a prompt (question 1), a monetary reward (question 2), a platform reward linked to gamification (question 3) and at last a curiosity trigger, already employed by Airbnb (question 4).

To ensure that the answers on these questions were accurate, a last section was added to the experiment that reflects on the intention to review questions. The purpose of this section was to verify that the information given in the intention to review section was accurate and usable. Due to the abstract context of the previously provided scenario (imagining the use of Airbnb), it was important to verify that a participant had no issues with imagining this context, and was familiar with Airbnb. Participants were presented with the following text:

> On the previous page I asked you to imagine a situation where you were using Airbnb. The provided context bombarded you with "Intention to review"-triggers. However, an important drive for how you act online is seeing the benefit of your actions. If there is no added value for you, the user, why would you do things? That's why I want to ask you the question:

After reading this text, users were asked the following question: *"Was it hard to imagine the situation presented in the previous point, about intention to review?"*. After answering the question, participants were asked to elaborate on their choice.

5.2 Results

In this section, the results of the first experiment are discussed. The questions in the experiment were either related to credibility or to intention to review.

5.2.1 Credibility

The results on the impact of credibility features on participants were varied. During the implicit test, not all participants based their decision of their favorite coffee machine on the

credibility features presented in reviews. Instead, some were lead by the content of the reviews. For example, information about how long a certain reviewer has been using a product had a strong influence on the preferences of people. Overall, the coffee machine that was picked the most was number four (the machine and review can be seen in figure 12). The main reason for picking this machine was that the product was owned the longest (for 1.5 years), which resulted in the participants trusting the review better.

Concerning the results of the explicit credibility tests, a hypothesis was that a certain selection of credibility features would be universally liked or would show a strong relation (e.g. someone that values user feedback highly could also be fond of feedback from the seller). Participants were able to rate the importance of credibility features on a scale of 1 to 5 (ranging from unimportant to very important). The means of the acquired values are presented in table 3 below.

Feature	Mean	Std. Dev.
Full name	2.793103	1.176536
Nickname	2.068966	1.162849
Social Media	2.620690	1.293218
Avatar	3.206897	1.081643
Verified	3.655172	1.009804
Endorsed	3.586207	0.982607
User feedback	3.034483	1.238742
Seller Response	3.000000	1.333333
Picture	3.137931	1.274030

Table 3: All tested credibility feature with their average score and standard deviation

The highest scoring features on average were the *verified user tag* and the *endorsement badge.* The score itself is not very high (in between moderately important and important), but it is interesting that on average, participants think that users acknowledged by a platform are more credible than other users. After analyzing the acquired data of the explicit credibility test, it became apparent that there was no strong correlation present in the test group. Most participants had different preferences throughout, with different reasons to feel that way. Making a correlation diagram of the explicit credibility data gives the following information:
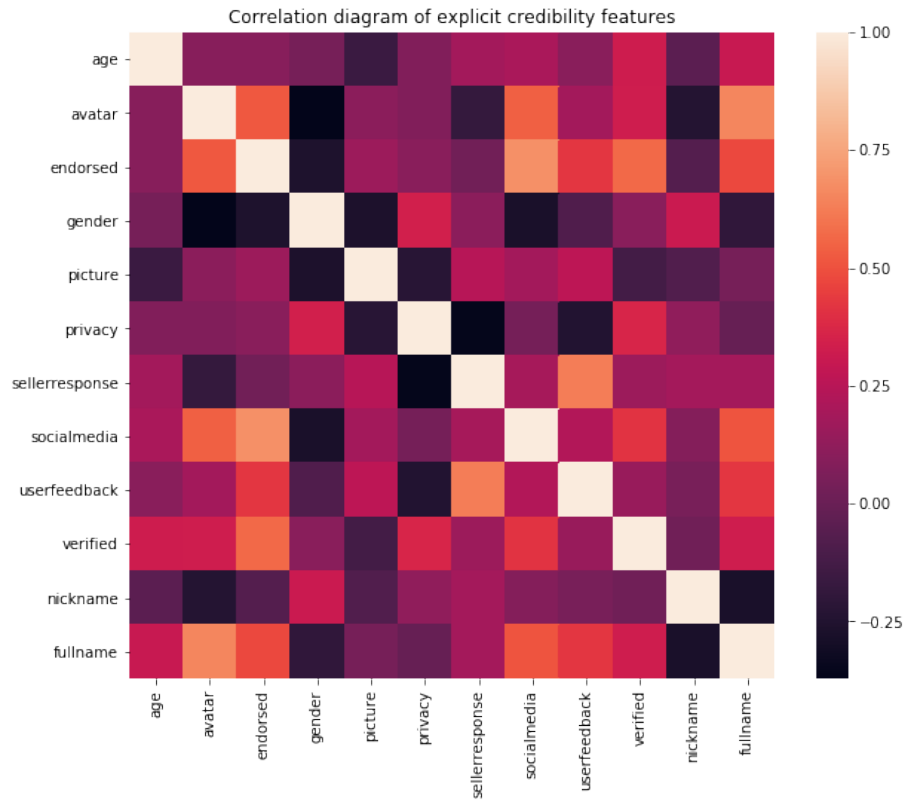
Figure 13: Correlation diagram of the explicit credibility test results of all 29 participants.

The rows and columns in the correlation diagram of figure 13 each relate to one of the tested credibility features. As can be seen in the color indication next to the diagram, there are no strong correlations visible over the whole group of participants. The black squares show a negative correlation of -0.3, which is insignificant. To see if there were interesting differences between different groups of participants, two divisions of groups were made. The first division is by age, separating all users that are older than 50 from the participants that were younger than 50. The resulting correlation diagrams can be seen in figure 14 below.
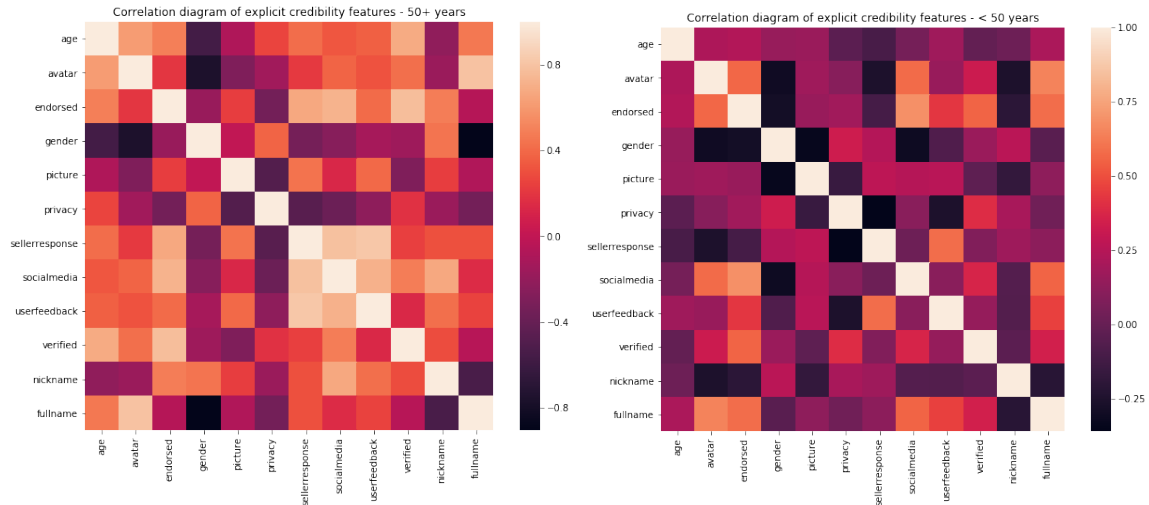
Figure 14: Correlation diagrams of the explicit credibility test results of the different age groups. Left: older than 50, Right: younger than 50

In this comparison, it is apparent that there are some significant correlations found in the group of older participants. Within this group, the people that said that feedback from other buyers is important for the credibility of a review, felt the same way about added feedback from sellers. Additionally, the same people felt that a link to someones social media profile in a review is important for its credibility. Within the younger group of people the correlations are not as significant, comparable to the correlation in results of all 29 participants as a whole. This visible difference could also occur due to the group of older participants being significantly smaller (eight participants) than the younger group (21 participants).

The second division is by gender, separating all male and female users. The resulting correlation diagrams can be seen in figure 15 below. This group is more evenly distributed in participants than the age groups, which results in a lack of apparent correlations. The significant correlation that can be seen is in the male diagram: men who think that personal avatars are important for credibility also feel that endorsed users are credible. This relation appears to be somewhat random, and should be confirmed by a study with more participants.
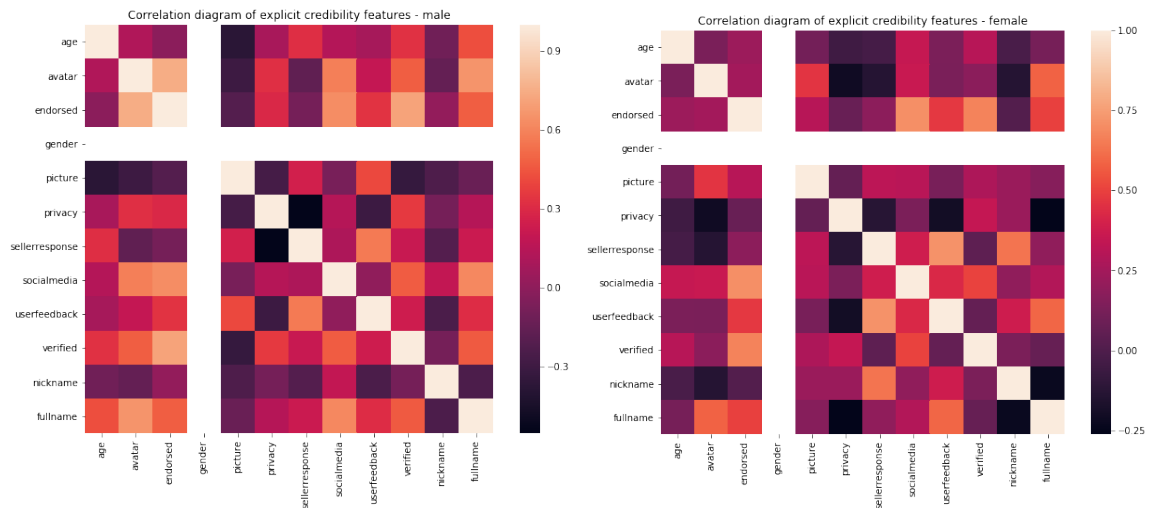
Figure 15: Correlation diagrams of the explicit credibility test results of the different gender groups. Left: male, Right: female

Questions about privacy also brought new results. People claim to value their privacy online very high: an average score of 4.142857 (in between important and very important). In the explanation of their choice, some insights were given. Some people prefer not to share their own personal information, but do find people that share their personal information more credible.

5.2.2 Intention to review

The results on the impact of intention to review triggers were varied as well. Quite some people (24%) explicitly explained that they don't feel the need to review something if a review score matches their expectation, or if other users already gave feedback similar to their own. This can be connected to the U-shaped reviews phenomenon: users tend to only review either very positive or very negative experiences, and leave the 'middle'-experience (which is what happens if a product matches their expectation) without feedback.

The results also showed that curiosity can be a strong trigger for people to leave a review. The question *"(...) Airbnb sends you a notification that lets you know that your host has left a review for you. You can only see what your host wrote about you if you leave a review too! Would this trigger you to leave a review?"* was answered positive by 13 participants (44%).

A simple explanation for their intention to review was given by four participants, who explained that they only leave a review if the hurdle to give their feedback is lower than the benefit gained from it. This sounds obvious, but this creates a clear picture of how intention to review works: the hurdle to leave a review should be low for users, while the gained benefit should be apparent for them.

Users were asked to reflect on the effectiveness of the provided Airbnb context. The provided context was succesful, as the majority of the participants (82.75%) were familiar

with the use of Airbnb.

At last, the question *"(...) Airbnb offers you platform credits that show on your profile after leaving a review. These credits could earn you different badges and platform privileges (E.g. more accomodation options in the future). Would these credits trigger you to leave a review?"* received mixed answers. People that frequently use a platform were able to see the benefits of this type of reward, which created a distinct difference in replies between platform users and non-platform users. People that do not frequently use online platforms experienced this gamification technique as something unpleasant. While theory (and online practice) suggests the effectiveness of gamification, the negative replies that were obtained told otherwise.

5.2.3 Distinction between platform users

One particular user gave detailed feedback from a P2P platform user perspective. This person (person A) explained that he frequently used Airbnb, and was very familiar with the provided scenario from section 5.1.7. Multiple people (24%), including person A, claimed that their intention to review was triggered by small rewards. The example rewards used in the experiment were either monetary (a discount) or credits that could be used on a platform (gamification).

Person A elaborated on systems that are used by Airbnb to encourage users to leave reviews in their reply. An example of a system was explained by person A: *"since a user wants to build a reputation (by having multiple positive reviews on their profile), it pays off to leave a review yourself. When other users come to your profile, they see that you're trustworthy"*.

The difference in reactions on credibility and intention to review triggers given by platform and non-platform users created an interesting contrast that needed to be researched further in the second iteration.

5.3 Discussion

The first iteration of the test showed interesting results. In this section, the test procedures and results of the first iteration are discussed, to aid in the design of the second user test.

5.3.1 Method of experiment 1

The most interesting findings in the first iteration were discovered in the answers to qualitative questions, which encouraged a focus on qualitative questions in the second iteration. The online context of the test proved to be effective, 83% of the participants were familiar with the usage of online platforms. The test was easily shareable online, and peers were able to provide extra help with forwarding the test online. At last, all the features that were tested in the first iteration originated from different features of the conceptual framework. Since there were interesting findings that were not included in the framework, the second iteration did not use the framework as a base, intending to expand on the already listed features.

5.3.2 Summary of interesting findings

The results of the implicit credibility test showed that the credibility of a review was influenced by the added credibility features, but also by the content of the review. What specific parts of content were perceived as credible is not certain, however.

In the results of the intention to review test, it became apparent that the U-shaped review phenomenon occurs in 24% of the participants. People do not want to leave their feedback if an experience or product is average, or if people already left a review that shares their own opinion.

Another interesting finding was that both credibility- and intention to review features are experienced differently by people that regularly use online platforms. People that claimed to be frequent users of Airbnb were more open to certain intention to review features. The people that mentioned their platform usage were not asked about it explicitly, so there were probably more people that frequently used online platforms that did not clarify this.

At last, an interesting find was detected during the qualitative analysis of the intention to review results. A presented reward system to trigger people to leave a review was experienced as bothersome by multiple people. This had to do with the gamification nature of the presented setup. While literature claims that people are incentified to engage with a platform if gamification systems are present, multiple participants felt that they were being deceived and withheld from certain parts of the platform.

5.3.3 Set-up changes

The changes in the set-up of the second experiment are based on the interesting findings from the first experiment. Four questions were presented that warranted extra research in the second iteration of the experiment:

1. In the first experiment, people were not asked about their platform usage. How can a better distinction between platform users and non-platform users be made?

2. Credibility features influence the credibility of a review, but the content influences the credibility too. What parts of the content and why influence review credibility?

3. Why do U-shaped reviews occur, and can this phenomenon be prevented?

4. Why is gamification experienced as a bad thing, and how can it be improved for users?

Since the set-up of the first test was successful, a similar survey-like test was warranted for the second iteration. The listed topics could be tested in a similar fashion as the first iteration, which presented the topics in multiple chapters. Thanks to the interesting results in the first experiment, the second experiment was focused on qualitative rather than quantitative questions.

6 Experiment 2 - Feedback driven

The second iteration of the user experiment had an exploratory nature, similar to the first iteration. Thanks to the results of the first experiment, more in depth research could be performed than in the first iteration. The second experiment could zoom in on certain interesting findings that were not initially part of the conceptual framework. The majority of the set-up is similar to the set-up of the first experiment.

6.1 Method

Just like in the first iteration, the method section is divided in a few key parts: Participants, Materials, Design and Procedure. When these parts have been discussed, the different segments of the experiment are explained in greater detail. Due to the similar set-up of the experiment compared to the first iteration, there will be many similarities in the method section.

6.1.1 Participants

People that participated in the first experiment were able to participate in the second experiment.

In the second experiment, 27 people participated. Of these people, there were 13 males and 14 females. To be able to compare the results of the first experiment to the result of the second experiment, the choice of demographic and the selection process were identical to the first experiment. Participants were contacted via Whatsapp, Facebook and by asking people in public university spaces to participate. All participants received the same test, there were no alternative versions or variables. The wide selection of participants allows to show insights and differences between demographics. The age distribution of the participants was similar to the first experiment, with a cluster of participants in their mid-twenties and a smaller cluster in their mid-fifties. The age distribution can be seen in figure 16.
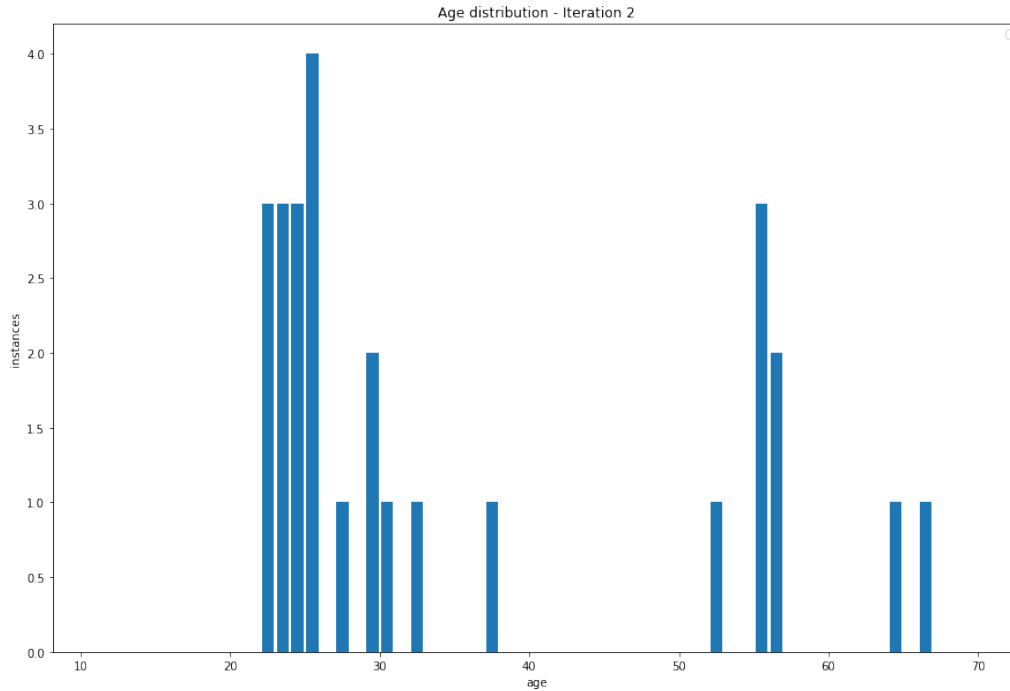
Figure 16: Age distribution plot of the participants of experiment 2

Important to note is that participants of the first experiment were able to participate in the second experiment. Due to the anonymous nature of the experiment, it was not possible to identify the users that participated in both experiments.

6.1.2 Materials

The used materials were almost the same as in the first iteration, as can be found in section 5.1.2. Instead of a selection of coffee machines, in this experiment a pair of wireless headphones were used as a review target. The pictures that were presented as stimuli were obtained from Tweakers.net. Tobón et al. (2020) described various academic gamification researches, which were used to determine the different gamification stimuli that were used in this experiment.

6.1.3 Study Design

Just like in the first iteration, this experiment was designed with exploration in mind. Interesting findings of the first iteration were used to base the second research on. Since the qualitative results of the first experiment were more valuable than the quantitative ones, this experiment will mainly comprise of qualitative questions. The main dependent variables remained the same: intention to review and credibility. The independent variables are the following:

1. Being a platform user (although this is not something that can be influenced by the research, it has an impact on intention to review and credibility)

2. Content of a presented review

3. Provided perspective on U-shaped reviews

4. Gamification techniques

Similar to the first iteration, other independent variables were age and gender. These variables were not the focus of the research, but they had the potential to aid in detecting interesting findings (and the age of participants has to be older than 16 due to the ethical committee requirements).

The controlled variables were the browser environment, the order of the questions during the experiment and the information that was presented during the test.

Similar to the first experiment, the second experiment used within-groups design. Just like in the first experiment, the reason for this was that by exposing all participants to all the triggers, individual difference would not be able to distort the results. The limited amount of participants would also be compensated by the use of within-groups design. In the section of the experiment that presents different reviews, the content of the reviews was randomly generated based on a selection of 'review sentences', to effectively detect credibility influencing content. The generated reviews were different for all participants, although the selection of sentences that was used to generate the reviews will be the same for everyone. More information about this can be found in section 6.1.6.

All participants received the same triggers, and were able to express their reactions that acted upon these triggers. The reason for using within-groups design was that by exposing all participants to all different triggers, individual differences would not distort the results. If between-subjects design was used, there would be a possibility that there were differences between the groups that could impact the results of the experiment. On top of that, having within-groups design proves to be effective with small groups of participants.

6.1.4 Procedures

The used procedures were almost the same as in the first experiment, as described in section 5.1.4. Exception is the provided study information which is slightly different in the second experiment, which can be read here:

Hello, it's me again, Floris! For new people: I am researching the behaviour of people in an online shopping environment. This is my second research, which is based on interesting findings of my first test. If you helped me out with the first test, thank you so much! This test will be shorter than the first one. Just like in the previous test, the obtained data will be anonymous, the only personal data needed is your age (you must be 16+ years old to participate) and your gender.

If at any point during the user test you wish to stop participating, you can just close the web page. You don't have to give a reason for this.

There are no risks in participating, this research project has been reviewed and approved by the EEMCS Ethics Committee. If any questions arise, feel free to contact me (Floris Veldhuizen, at f.l.m.veldhuizen@student.utwente.nl) or my supervisor (Dennis Reidsma, at d.reidsma@utwente.nl). Finally, you could also reach the Ethics Committee at ethics-comm-ewi@utwente.nl.

In the next sections, the different parts of the experiment are explained in more detail.

6.1.5 Identifying platform users

The second experiment started with the identification of (P2P) platform users. As was discussed in section 5.2.3, P2P platform users appeared to respond differently to credibility and intention to review features than non-platform users. In the first experiment, there was no way to tell if participants regularly used P2P platforms, except if they told about their platform usage on their own initiative. To make the distinction between platform users and non-platform users clearer, this experiment explicitly asked users about their platform usage. The participants were asked about their platform usage after being presented with the following text:

After analyzing the responses from my previous test, it became clear that people who are regular users of online platforms respond differently to reviews than people who are not. When I am talking about online platforms, I refer to platforms like Airbnb or Uber: platforms where people (buyers) can pay for services from other users (sellers). This question will make it easier for me to determine the differences between platform users and non-platform users in the next part of my research.

Following this explanation, participants were asked *"Do you use online platforms?"* to which they could answer with the following options: "Never", "Sometimes" and "Often". After participants answered this question, they were asked to elaborate on their platform usage with the following question: *"What platforms do you use, and how often?"* (unless they answered "Never"). In the end, this section was designed to aid in understanding what the needs of a platform user are compared to the needs of a non-platform user.

6.1.6 Content specific credibility

Similar to the coffee machine test found in section 5.1.5, this section of the experiment presented multiple reviews to participants. However, rather than testing the effect of credibility fea-

tures on participants, this experiment focused on the influence of the content of a review on its credibility. Instead of a selection of six different products with reviews, this experiment had one product with six reviews. The reason for using only one product was that it was impossible to have reviews that were in line with each other if they all review another product. The product that was reviewed is a pair of wireless headphones (the Sony WH-XB900N), due to the easy access to product pictures and the versatility of the product (there are many technical details that can be talked about in a review). The presented reviews were content focused, with only a basic selection of credibility features (as seen in figure 17). The reviews are generated based on a selection of specific sentences that were combined to form a review. Below, the context provided to the users is shown:

> You are looking for a new pair of wireless headphones. After looking around, you stumble upon this pair that seems pretty good (the Sony WH-XB900N). You have enough money to buy it, and now want to see what other people think about these headphones. After looking for online reviews, you find the reviews displayed below. What review would influence you to buy (or not buy) this pair of headphones? Why? *You can click on the review that feels the most relevant.*

After this text, the participants were presented with six randomly generated reviews. All the reviews contain a similar ordered structure:

1. The name of the reviewer is shown, accompanied by a product rating and an anonymous avatar.

2. The review starts with a time frame, how long has the reviewer owned the product?

3. The reviewer provides (technical) details about the product.

4. The reviewer describes their relation with the brand or technology.

5. The reviewer finishes the review with a personal note.

6. (Optional) Two reviewers provide a picture with their review.

Every part of a generated review was randomly selected out of premade sentences that fitted the structure. The reviews were randomly generated within this structure to find out what specific parts of review content influenced people and generated credibility. For example, do people trust a review more if the reviewer claims to be a technical expert? Since the content of the reviews varied per participant, it was possible to isolate specific parts of content that were a strong credibility trigger for multiple people. To prevent an inconsistency within the reviews, some combinations of sentences were impossible to be generated. All the used sentences and their respective category are listed below:

Full name

1. Johan de Vadder

2. Ronald Jonkheer

3. Bart Ochtendschalker

4. Geert Snoeibrug

5. Marian Ketelburgh

6. Eva van Kampen

Timeframe

1. 3 days ago I got my hands on these headphones.

2. I got my fresh pair of Sony headphones last week.

3. One month ago I decided to get new headphones.

4. I've had this product for about six months.

5. Last year I purchased them, and I have used it over the year.

6. I bought my pair 1.5 years ago.

Details

1. The headphones are very durable, the material feels strong. When I dropped them, there wasn't a single scratch!

2. The rich deep bass works really well for upbeat tracks, like I'm standing right in the club.

3. It feels very comfortable on my head, I can wear it all day long.

4. It looks amazing and finishes a few of my outfits. Its neutral color makes it easy to match with everything!

5. I love how balanced the sound of the headphones is, a very natural sound that fits all music.

6. The battery on these headphones is amazing, they last one week of daily use on a single charge.

7. The noise cancelling feature is great, living next to a cafe makes me appreciate the clean music even more!

Negative details (Only used in the negative review)

1. I felt that the range is too focused on the bass, which makes regular songs sounds very muddy.

2. While the material is supposed to be durable, after carrying it in my bag I could already see minor damages.

3. I think they look way better on the online pictures than in real life. It feels like a scam.

4. The noise cancelling feature works abysmal, I can still hear noises around the house easily through the music.

5. After wearing them all day my head hurts, they don't feel comfortable after a few hours of wearing...

Relation to the brand or technology

1. I have bought other devices from the same brand, they never disappoint.

2. I worked in audiotech for a few years, and the guys that are working on this clearly know what they are doing.

3. This is the first time that I buy headphones from this brand, but I am sure it won't be the last.

4. After looking at other options online, this one seemed to be the best value for money.

5. Before I got this pair of headphones, I have never bought anything from this brand before.

6. I bought something from this brand before, but was not particularly happy with it. I still wanted to give them a chance with these headphones, though.

Personal note

1. I enjoy listening to my favorite songs with them!

2. I would recommend this product.

3. I would buy this again if it ever breaks.

4. This is probably my new favorite pair of headphones!

5. They do what they must do, which is good.

6. My experience with it has been bad, I'm not happy with my new headphones. I wish I got other ones.

Five of the six generated reviews were positive, while one review was negative. This review was intentionally not in line with the other reviews. The negative review was added to see if people tend to pick the review that was not in line with the other reviews to influence their decision. Two of the generated reviews had a picture, to find out if the added picture made a difference to the credibility of the reviews. There was a sentence added to these reviews that mentioned the picture, to make the review feel more natural to participants. Finally, there was a selection of negative details that was used to create a realistic negative review. An example of a generated review is presented in the figure below:
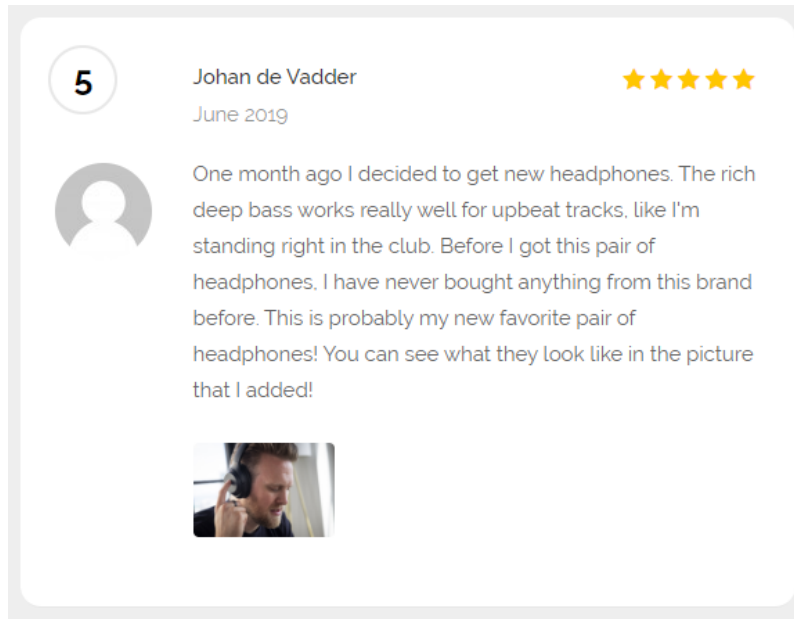

Figure 17: Example review in iteration 2 of the experiment.

6.1.7 U-shaped feedback

The results of the first experiment indicated that people were hesitant to review experiences that are not great, but also not bad. This phenomenon is called U-shaped feedback, and is already briefly discussed in section 2.2.2. To understand why this phenomenon happens and if it can be prevented, this section of the experiment was designed. Users were presented with the following paragraph:

> When leaving feedback online, a lot of people don't feel the need to review the 'middle'-experience, which results in products that get either very negative or very positive feedback (hence, the U-shape). While this doesn't seem like a big problem, as the average rating might still be accurate, it means that a lot of opinions are not heard. For many people, it is important that there are enough reviews on their products. They need reviews to build trust in their product and to learn how to improve their service. If a review is in-line with existing reviews, it still contributes to the reliability of the feedback. It allows sellers to prioritize different problems that come up in similar reviews.

Followed by this text, users were asked the following question: *"If you knew how important your feedback was for other people, would you review a 'middle'-experience?"*. The goal of this question was to find out if participants could be motivated to leave a review of a 'middle'-experience, and to understand their motivation behind not leaving a review in such a situation. To further investigate the motivations of the participants concerning their feedback, the following text was presented:

> Giving feedback can be made easier by giving rating options that require little effort. Think about implementations such as a like/dislike button, a 5-star rating system or pressing an 'emotion'-button. Because Youtube got so many U-shaped feedback on their videos, they implemented a thumbs up / down button and abandoned their star-rating system.

This text was followed up with the question: *"Would a simple rating system trigger you to leave feedback? What does your ideal rating system look like?"*. The question was designed to gain a better insight in reputation systems that are preferred by participants. Finally, the last part of this section was used to test a theory that I created after analyzing the results of the first usertest. The theory was presented to participants with the following text:

> In my first experiment I noticed something interesting, shared by multiple people that use online platforms or order stuff online. When they think that a product rating is fair, or if they read a review where people express an opinion similar to theirs, they don't feel the need to leave their own feedback. Based on the number of product sales versus the number of reviews, you can make an estimation of the number of people that agree with a current rating, but decide not to give a rating themselves. To find out if this is accurate, I want to ask you the question:

After reading this, users were presented with the question: *"If you buy something online and you think that the product rating is fair, or if there is already a review with your opinion, would you still write a review?"*. Answers on this question could lead to a better understanding of online ratings.

6.1.8 Gamification

Contrary to what literature suggested, the results of the first experiment indicated a negative perception of gamification techniques that were used. Gamification was mentioned in section 3.6 and is utilized by companies such as Airbnb (see section 3.3). The results of the first experiment lead to the design of this section of the experiment. Why was gamification experienced as a bad thing, and how can it be improved for users? Users were presented with different gamification techniques, based on a gamification theory by Tobón et al. (2020). Their paper suggests that the influence of gamification can be explained by the mechanism of reward, challenge, meaningful, and interactivity. These four categories formed the base of the different gamification scenarios that were presented to users, to find out what types of gamification are considered pleasant, and what types are not. At the start of this section, users are presented with the following text:

> Gamification is when elements of playing a game (E.g. rewards, experience, competition) get applied to other areas of activity. It is often used to make simple tasks more interesting and rewarding, and can also be used as a marketing technique to engage customers with a product or service. In my first experiment I noticed that multiple people do not like gamification at all. While both theory and literature support the effectiveness of gamification, I wanted to see what you think about gamification, and if it is effective at all.
> On this page, I am proposing a few versions of gamification that can be used on online platforms. Would the techniques help to make a platform more fun, or help you engage with a platform? Let me know what you think.

The following questions with relation to the aforementioned four categories are presented to the participants:

1. (Reward) - Exceptional platform users get rewarded with platform benefits and a visible badge (think about Airbnb's superhost badge). *Would this trigger you to engage with the platform? Explain why!*

2. (Challenge) - The writer of the best review of a product will get the next edition of the product for free. *Would a challenge like this trigger you to write a review? Explain why!*

3. (Meaningful) - Everytime a product receives a well written review, a tree will be planted by the platform (to save the planet)! *Would you help out by writing a review? Explain why!*

4. (Interactivity) - The platform has a lot of personal customization options. You can rate other reviews and product features, and customize your public profile. *Would this trigger you to engage more with the platform? Explain why!*

6.2 Results

In this section, the results of the second experiment are discussed. The questions in the experiment were related to credibility and intention to review.

6.2.1 Platform users

Most of the participants (21 people) are occasional platform users, although not many claimed to use online platforms often (only 6 participants did). The most popular online platform that was used is Airbnb, which was used by 17 people. A close second is Marktplaats, with 15 people. The third most used platform with 6 people is Thuisbezorgd.nl.

6.2.2 Credibility

This section describes the results acquired from the test described in section 6.1.6. It was interesting to note that none of the participants chose the negative review (the review that was not in-line with the other generated reviews) as being the most influential. Three participants did mention the negative review and told that it did make them consider the negatives of the product. However, in all cases, the 5 positive reviews canceled out the single negative review. To elaborate on the impact of negative reviews, one of the participants told that they analyze reviews by looking up the most negative reviews of a product and then comparing them to the positive ones.

Overall, the most credible factor in the generated reviews was the authority of an expert. 13 participants stated that the feedback of an expert on the topic was what influenced them to pick a review. Building on this, four people mentioned that the type of language that was used in a review was the most important factor for credibility to them. Eight participants mentioned that it was important for them to be able to compare the quality of different products, which related to reviews that mentioned the reviewers relation to other products of a brand.

Another factor that added credibility for many people was the time that a reviewer owned a product. The most credible time was also the longest (1.5 years), after that came ownership of one year, and then six months. What was interesting to see was that pictures did not seem to influence credibility for the majority of the participants. There were 11 participants that picked a review with an added picture, however, only three of these participants mentioned that they preferred them because of the picture.

Additionally, the gender of the reviewer did seem to have an influence on the participants' sense of credibility. The majority of the picked reviews (63%) were written by men. The gender of the participant had no influence in this preference, both male and female participants shared a preference for male reviewers.

Other interesting findings were that a participant said that the total amount of reviews of a product greatly influenced their decision to buy a product. This is in line with another participant's opinion: both were looking for consistency in reviews to base their decisions on. It appears that being a platform user does not enhance a specific credibility trigger compared to non-platform users, but this might also have to do with the low amount of non-platform users to compare the results with.

6.2.3 U-shaped feedback

This section describes the results from the test in section 6.1.7. Participants were informed about the importance of reviews, even if it concerns feedback that reviews the 'middle'-experience. After being informed, participants were asked if they would provide a review of an average experience. Out of the participants 10 people would not, 8 people had conditions or doubts and 7 said that this would convince them to leave a review.

As the second question of this section, participants were asked what they thought was the ideal rating system for them. The most common feedback on this question was that their preferred rating system is quick and descriptive. If giving feedback required one click or tap, people didn't feel hesitant to leave their quick review. Ten participants explicitly mentioned that the time spent on giving a review is the determining factor for whether they decide to do it or not. This philosophy goes hand in hand with the feedback of the six participants that just want an easy rating system (e.g. a like/dislike rating system). Six participants claimed that they were content with a 5-star rating system, a system that is commonly used online. The preference for this rating system can be linked to familiarity with a system like this. Four participants said that while they felt that a 5-star rating system is convenient, it lacks depth. A proposed improvement that was suggested was a technique that is already adapted by Airbnb: rating the different aspects of a service (like Airbnb does with cleanliness, good contact with the owner, etc.). Two participants mentioned emotion rating systems, where you can press on a smiley face to indicate your opinion or emotions. One user suggested a rating system with 4 stars, so there could be no neutral option. Interesting to note is that some of the mentioned rating systems have been discussed in the State of the Art section, as seen in figure 6.

Finally, participants were asked if they would leave a rating online if they felt that a product rating is already fair or when another reviewer expressed a similar opinion to theirs. As was expected, the majority of participants (19 people) answered no. Within this majority were all 6 participants that are non-platform users. Six participants mentioned that they would leave feedback if there was an option to agree with already given reviews (e.g. Amazon's "X people found this review helpful"). Six people said that as long as their feedback would be easy to give, they would not be bothered by leaving a rating. Finally, 2 people claimed that they would always leave a rating online even if they felt that their opinion was already shared. One of these was also a frequent platform user.

6.2.4 Gamification

This section describes the results of the last section of the experiment, section 6.1.8. In this section, effects of different gamification techniques were tested. People were asked if using certain gamification techniques would trigger them into leaving feedback. Results were varied between utilized techniques:

Reward based: The participants were confronted with a reward based gamification technique via the following question: *"Exceptional platform users get rewarded with platform benefits and a visible badge (think about Airbnb's superhost badge). Would this trigger you to engage with the platform? Explain why!"*. The response to this question was highly dependent on the relation of the participant to a platform, and the added benefits of the gained reward. If the benefits gained from a reward fit their personal desires, 14 people claimed that they would

leave a review if they would receive a reward for leaving their feedback. Of these 14 people, six did not mention these conditions and outright said yes. However, 13 participants claimed not to be sensitive for this kind of gamification technique.

Challenge based: This gamification technique was presented with this question: *"The writer of the best review of a product will get the next edition of the product for free. Would a challenge like this trigger you to write a review? Explain why!"*. An interesting recurring response to this question was that many people were doubting the integrity of provided feedback if it was stimulated by a challenge. Why would a platform pick a negative review as the best, and who decides what the best review is? What would the integrity be of the winning review? Even though these doubts existed, eight people said that this would trigger them to write a review. Six participants were not sure, and 12 participants claimed to not be influenced by a challenge as was presented.

Meaningful based: The meaningful gamification technique was shown via this question: *"Everytime a product receives a well written review, a tree will be planted by the platform (to save the planet)! Would you help out by writing a review? Explain why!"*. Results showed that this was the most effective technique, with 19 participants that would help out by writing a review. Of the willing participants, some people were afraid of possible abuse and also wanted to see a clear overview of planting progress / conditions. There were more sceptical people, 5 participants would not help and did not believe that this could be feasible for a company.

Interactivity based: The final question of this section was an interactivity technique, asked about via this question: *"The platform has a lot of personal customization options. You can rate other reviews and product features, and customize your public profile. Would this trigger you to engage more with the platform? Explain why!"*. This technique proved to be least effective, with only five participants that would be more likely to engage with a platform that uses this technique. Of these five people, all of them were mostly interested in specific interactivity aspects, like the ability to rate other reviews or the ability to rate specific product features. Two participants would only consider this technique as effective if they were invested in a platform. The other 19 participants were not stimulated by interactivity.

Interesting to note was that on average none of these techniques had an impact on participants that said that they did not use platforms. Only two of the six platform users responded positively to any of the presented techniques.

7 Discussion

7.1 Result analysis

This section discusses the combined results of the literature research, the state of the art research and the two iterations of the performed experiment.

7.1.1 General findings

It was hard to detect differences in online behavior between platform users and non-platform users. The clearest difference was found in the way that participants reacted to gamification triggers, with the majority of platform users being more sensitive to gamification techniques than the non-platform users.

7.1.2 Contradictions

Results that contradicted with theories from the literature were found in the first experiment, when a gamification technique was tested and multiple participants reacted negatively to it. To confirm if gamification techniques were experienced negatively in general, multiple variations of gamification techniques were tested in the second experiment. It appeared that different gamification techniques work for different people. In three out of four gamification methods, at least 44% of the participants were not comfortable with these techniques being used. The gamification that was tested in both of the experiments was presented to the participants in an explicit way. This could have influenced the way that people react to them.

7.1.3 Confirmations

When looking at online reviews, the majority of people look for reviews of an authority on the subject to base their own opinions on. This behavior was also mentioned by Jessen and Jørgensen, as can be seen in section 2.3.3. To add to that, the period that a product is owned before a review is written also influences the level of credibility of a review. This is a logical consequence, as being experienced with a product results in being more familiar with its weaknesses and strengths, which creates an expert opinion.

7.2 Limitations of the study

The small scale of the study made it difficult to detect more than trends in the acquired data. Especially the small group of people that did not use P2P platforms was hard to compare to the platform user group, as there were only 6 participants that never use online platforms. Another limitation of the study that was apparent was the conceptual nature of the set-up. Participants can claim to behave in a certain way while acting online, but this can differ from how they would really behave. Having access to more participants that interact with reputation system in a natural way would increase the integrity of the results.

Additionally, participants of the first study were also able to participate in the second study. The fact that this could influence the results of the second experiment was not considered

before the experiment was distributed. Due to the participants being anonymous, it was not possible to detect who of the participants participated in both of the studies. While the impact of the first study on the second study is not clear, there was a period of 40 days in between both studies. While credibility and intention to review were the main topic in both experiments, the questions that were asked in the second experiment were aiming at specific reputation system features that were not highlighted in the first experiment.

7.3 What features in reputation systems have the strongest influence on their users?

To conclude the discussion section, an answer to the following questions is provided:

1. What credibility features in reputation systems have the strongest influence on its users?

2. What intention to review features in reputation systems have the strongest influence its users?

While the amount of data that is needed to reliably answer these questions was not available due to the scale of the experiment, the acquired data was able to indicate trends.

It appears that the most influential credibility features of a reputation system can be divided in content specific features and in reputation system features. On average, the highest scoring credibility reputation system features during the first experiment were the *verified user tag* and the *endorsement badge*. This indicates that people prefer feedback given by authorities, or by users that are confirmed to be authentic by a platform. This confirms the theory of Jessen and Jørgensen, as discussed in section 2.3.3. Another recurring aspect of eliciting credibility in reviews was the addition of a picture to a review. While the participants did not mention the picture, 40% of participants of the second experiment felt that a review with a picture added was the most influential to them. In terms of content specific features, the time that a product was owned by its reviewer and the expertise of said reviewer influenced the decision of the participants the most. This, again, confirms the aforementioned theory of Jessen and Jørgensen. To add to the content specific features, there appeared to be a slight preference (by 63% of participants) for the opinion of a male reviewer as opposed to the opinion of a female reviewer. This might have to do with the technical nature of the tested product review, and the bias that men are more knowledgeable in this regard. It seems that if a reputation system is able to show the authority and the expertise of its users, people tend to find the provided feedback the most credible. Product ratings that are formed by many reviews together also form a credible source for people.

Most of these influential credibility features are dependent on the people that leave their feedback. This is where intention to review becomes important: if credibility is gained by the quantity of the reviews and the authority of the reviewer, a platform should stimulate its users to leave their feedback. This moves us to the second question: *"What intention to review features in reputation systems have the strongest influence its users?"*. A strong intention to review trigger related to gamification was if something meaningful was linked to the giving of their feedback. 70% of participants claimed that this would convince them to leave a review. In the experiment, meaningfulness was represented by planting a tree for every review, but this could be a costly solution.

Another gamification related trigger was gaining platform benefits by leaving feedback. Reward based gamification proved to be an effective intention to review trigger as long as the person that gains platform rewards is an active user of the platform. An overarching decentralized reputation system that is used by multiple platforms could show a potential outcome, as platform benefits could be gained over the different platforms that would implement said system. If someone would be able to gain benefits for their preferred platform while using the reputation system of another platform, there would definitely be more intention to review. This could be an interesting thought for future work.

An alternative way to elicit intention to review is by implementing a very basic reputation system. If giving a rating only costs one click to a user, most of them would not be hesitant to leave their feedback. Opinions on what the ideal reputation system should look like were varied, as ease of use correlates negatively with the depth of a reputation system. A solution that was mentioned aims for the middle ground: if a product has multiple categories that can be rated with an easy rating (in case of Uber think about politeness, travel time, cleanliness, etc.), there is both ease of use and depth to a system. To elaborate on this, people enjoy a quick option to validate other feedback instead of giving their own.

Finally, in a P2P platform context the two way double blind review system that is already used by Airbnb and Uber proves to be effective. 44% of the participants in the first experiment claimed that curiosity is a strong trigger to motivate them to leave their feedback.

8 Conclusion

8.1 Overall summary

The first obvious conclusion that can be derived from the results of both experiments is that all people had their individual preferences. Everyone values different reputation system features, there was not a single feature that was disliked or liked by all participants. That being said, there are detectable trends that indicate that some reputation systems features had more impact on people than others. The next important conclusion is that there are many different factors that influence the performance of a reputation system. Next to the credibility and intention to review features that can be added to a reputation system, the textual content of a review greatly influences its credibility. If a reviewer is an expert on the presented topic or if they owned a product for a long time, their review will be generally perceived as credible. Another factor that influences the performance of a reputation system is the environment that it exists in. A platform that has an established user base and has been around for a while is naturally more credible than a platform that is new and unknown. A reputation system can influence trust, but cannot replace the reputation of a platform as a whole. There is also the interesting balance between credibility and intention to review, as a very credible and in-depth system usually requires complex feedback to maintain this credibility. A reputation system that needs complex feedback will elicit less intention to review in its users, as a big factor that influences intention to review is the time that needs to be spent on the given feedback. Finding the correct balance between credibility and intention to review seems to be vital for the design of a successful reputation system.

8.2 Updated conceptual framework

This section aims to answer the research question *"How can the conceptual framework be refined so it can aid in reputation system design?"* An important part of improving the initial framework from chapter 4 was making it usable as a tool that can aid in reputation system design. By ranking certain features as effective, the framework can be used as guide that is able to indicate if a feature is missing, if a feature is effective, and why a feature is effective. After testing features of the initial framework in different experiments, features could be validated as effective or ineffective.

Based on all the performed experiments and the initial conceptual framework, finally an improved version of the conceptual framework was created. The improved framework includes new reputation system features and adds a "+"-mark to features that have proven to be effective during the experiment. Features that were perceived negatively or did not show much promise received a "-"-mark. Features that did not get a mark did not perform exceptional (getting some positive reactions, but not by a majority). Finally, features with a dashed border have not been tested. However, the untested features still have a place in the framework due to their proven effectiveness in the literature and state of the art studies.

The improved framework has color coded features:

- Blue: indicates a feature related to rating systems

- Red: indicates a feature related to P2P platforms

- Green: indicates a feature related to social media or social presence

- Orange: indicates a feature related to the categorizing and transparency of reviews

- Purple: indicates a feature related to gamification

- Grey: indicates a feature related to the content of a written review

- White: indicates a feature that is not limited to one of these categories

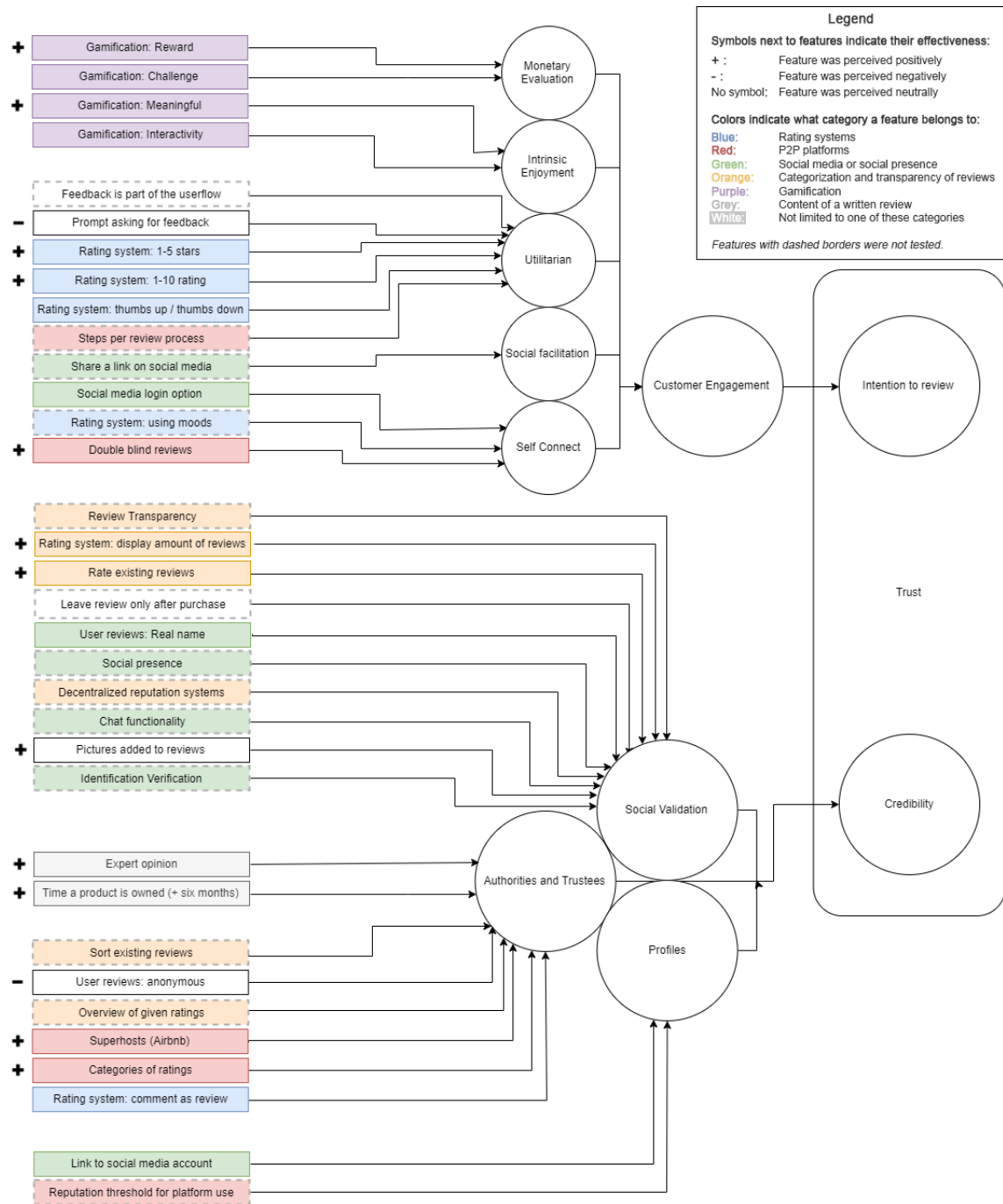The result of the improvements can be found in figure 18 below.

Figure 18: Updated version of the conceptual framework created in chapter 4.

8.3   Answering the research question

With an improved version of the conceptual framework, the main research question could be answered: *"How to design a good reputation system for an online peer-to-peer platform?"*.

To answer this question, the definition of a good reputation system had to be made. A literature study was performed to define the metrics that can effectively indicate the performance of a reputation system. Most of the relevant literature was focused on the perspective of the buyer, which is commonly the largest group of users on a platform that interact with reputation systems. The literature study resulted in the following definition of a good reputation system: *the performance of a reputation system can be defined by how credible the system is, and how effective it elicits intention to review in its users.* With these metrics in mind, a conceptual framework was designed that shows how a reputation system influences both credibility and intention to review on a feature level. After performing multiple experiments to validate and explore these features, an improved version of the conceptual framework was designed, as can be found in figure 18. While this improved framework is still conceptual and based on experiments with a small sample size, it is able to show a multitude of features that are used in reputation system design effectively. The framework categorizes the features in a structured way, and shows which features were received positively by the participants in the performed experiments. The improved framework can be used to design reputation systems or to analyze them. When a designer uses the framework to design a reputation system, they can pick features that were perceived positively or that relate to their specific desires, while avoiding the bad ones. The framework lists specific reputation system features that relate to P2P platforms, which makes it an effective tool to design a reputation system for online P2P platforms.

Finally, as was mentioned in section 8.1, the performance of a reputation system is influenced by the environment it exists in. A reputation system that includes effective features can still perform underwhelming if it is implemented on an untrustworthy platform.

8.4   Recommendations

In this thesis, the focus of the research and the performed experiments was on reputation systems on a feature level. These features influence the credibility and intention to review of a reputation system. In future research it could be interesting to find out if a reputation system can be designed and improved on something other than on the feature level. Credibility and intention to review branch of to multiple mediating factors (e.g. social validation, customer engagement) that are influenced by features, but these mediating factors themselves could also be used as a research topic.

Future research could also be used to take an in-depth look at the different perspectives of people that interact with reputation systems. This thesis mainly focused on reputation systems from the perspective of the buyers, but the perspective of selling users or the platform itself could yield different results or create new insights to further develop a good reputation system. Elaborating on the change in perspective, it could also be interesting to do a comparison study between reputation systems on P2P platforms and B2C platforms. For example, if a reputation system on a P2P platform and a B2C platform have the same reputation features, will people

react to them differently?

If further research is performed on credibility and intention to review features, it could be interesting to measure the performance of these features in a continuous way. Having continuous values allows the results to be more effectively analyzed with machine learning methods, which could detect patterns that are hard to detect otherwise. To further enhance machine learning techniques in future studies, more user data should be gathered. The most effective environment to test in would be an environment where participants are not aware that they are doing a test. Having access to an online P2P platform to perform tests with a big user base would be the ideal scenario.

In the results of the content specific credibility experiment, participants were shown 5 positive reviews with one negative review. While the negative review was mentioned multiple times, none of the participants picked the negative review as being the most influential. This is likely due to the one negative review feeling like an outlier. It would be interesting to find out what an acceptable ratio of negative versus positive feedback is, and where people draw the line.

In future studies, the limitations mentioned in section 7.2 should be taken into account. If similar studies would be performed, it would be wise to invest in gathering more participants to gain more significant results. This study is able to give direction to future reputation system design, but more data is needed to objectively base designs on.

References

M. Basili and M. A. Rossi. Platform-mediated reputation systems in the sharing economy and incentives to provide service quality: The case of ridesharing services. *Electronic Commerce Research and Applications*, 39:100835, 2020. ISSN 1567-4223. doi: https://doi.org/10.1016/j.elerap.2019.100835.

V. Blazevic, W. Hammedi, I. Garnefeld, R. Rust, T. Keiningham, T. Andreassen, N. Donthu, and W. Carl. Beyond traditional word-of-mouth: An expanded model of customer-driven influence. *Journal of Service Management*, 24, 06 2013. doi: 10.1108/09564231311327003.

D. Bulmer and V. DiMaurio. *The New Symbiosis of Professional Networks: Social Media's Impact on Business and Decision Making: Aresearch Study*. SNCR Press, 2010.

A. C. Calheiros, S. Moro, and P. Rita. Sentiment classification of consumer-generated online reviews using topic modeling. *Journal of Hospitality Marketing & Management*, 26(7):675–693, 2017. doi: 10.1080/19368623.2017.1310075.

A. Costa, J. Guerreiro, S. Moro, and R. Henriques. Unfolding the characteristics of incentivized online reviews. *Journal of Retailing and Consumer Services*, 47:272 – 281, 2019. ISSN 0969-6989. doi: https://doi.org/10.1016/j.jretconser.2018.12.006.

B. de Langhe, P. M. Fernbach, and D. R. Lichtenstein. Navigating by the Stars: Investigating the Actual and Perceived Validity of Online User Ratings. *Journal of Consumer Research*, 42 (6):817–833, 09 2015. ISSN 0093-5301. doi: 10.1093/jcr/ucv047.

C. Dellarocas. The digitization of word of mouth: Promise and challenges of online feedback mechanisms. *Management Science*, 49(10):1407–1424, 2003. doi: 10.1287/mnsc.49.10.1407.17308.

R. Dingledine, M. J. Freedman, and D. Molnar. Chapter 16: Accountability. *Peer-To-Peer: Harnessing the Power of Disruptive Technologies. O'Reilly and Associates, Inc*, 2001.

B. Fogg. *Persuasive Technology: Using Computers to Change What We Think and Do*. Morgan Kaufmann Publishers Inc., San Francisco, CA, USA, 2003. ISBN 9780080479941.

B. J. Fogg and H. Tseng. The elements of computer credibility. page 80–87, 1999. doi: 10.1145/302979.303001.

T. Grandison and M. Sloman. A survey of trust in internet applications. *IEEE Communications Surveys and Tutorials*, 3:2–16, 01 2000. doi: 10.1109/COMST.2000.5340804.

N. Hajli. Ethical environment in the online communities by information credibility: A social media perspective. *Journal of Business Ethics*, 149(4):799–810, Jun 2018. ISSN 1573-0697. doi: 10.1007/s10551-016-3036-7.

E. Hargittai, L. Fullerton, E. Menchen-Trevino, and K. Thomas. Trust online: Young adults' evaluation of web content. *International Journal of Communication*, 4(0), 2010. ISSN 1932-8036. URL https://ijoc.org/index.php/ijoc/article/view/636.

F. Hawlitschek, B. Notheisen, and T. Teubner. The limits of trust-free systems: A literature review on blockchain technology and trust in the sharing economy. *Electronic Commerce Research and Applications*, 29:50 – 63, 2018. ISSN 1567-4223. doi: https://doi.org/10.1016/j.elerap.2018.03.005.

J. Jessen and A. Jørgensen. Aggregated trustworthiness: Redefining online credibility through social validation. *First Monday*, 17(1), 2011. ISSN 13960466. doi: 10.5210/fm.v17i1.3731.

A. Jøsang, R. Ismail, and C. Boyd. A survey of trust and reputation systems for online service provision. *Decision Support Systems*, 43(2):618 – 644, 2007. ISSN 0167-9236. doi: 10.1016/j.dss.2005.05.019. Emerging Issues in Collaborative Commerce.

V. Katz. Regulating the sharing economy. *Berkeley Technology Law Journal*, 30(4):1067–1126, 2015. ISSN 10863818, 23804742.

R. D. Lankes. Trusting the internet: New approaches to credibility tools. *The John D. and Catherine T. MacArthur Foundation Series on Digital Media and Learning.*, 2008.

L. Liu and M. Munro. Systematic analysis of centralized online reputation systems. *Decision Support Systems*, 52(2):438 – 449, 2012. ISSN 0167-9236. doi: https://doi.org/10.1016/j.dss.2011.10.003.

A. Ljung and E. Wahlforss. *People, Profiles and Trust: on Interpersonal Trust in Web-mediated Social Spaces*. Trita-CSC-E. Lulu.com, 2008. ISBN 9781409229421. URL https://books.google.nl/books?id=myAAyj0hmq8C.

A. G. Mauri, R. Minazzi, M. Nieto-García, and G. Viglia. Humanize your business. the role of personal reputation in the sharing economy. *International Journal of Hospitality Management*, 73:36 – 43, 2018. ISSN 0278-4319. doi: https://doi.org/10.1016/j.ijhm.2018.01.017.

F. Mazzella, A. Sundararajan, V. Butt D'Espous, and M. Möhlmann. How digital trust powers the sharing economy. *IESE Insight*, Third Quarter:24–31, 09 2016. doi: 10.15581/002.ART-2887.

J. Meijerink and E. Schoenmakers. Explaining reputation inflation on airbnb: The role of users' perceived service quality. 2019. Reshaping Work 2019 ; Conference date: 24-10-2019 Through 25-10-2019.

S. M. Mudambi and D. Schuff. Research note: What makes a helpful online review? a study of customer reviews on amazon.com. *MIS Quarterly*, 34(1):185–200, 2010. ISSN 02767783.

W. I. O'Byrne. Facilitating critical thinking skills through content creation. 2009.

L. Pettingill. Trust without knowledge: How young persons carry out research on the internet. *GoodWork Project Report Series*, (48), 2006.

P. Resnick, K. Kuwabara, R. Zeckhauser, and E. Friedman. Reputation systems. *Commun. ACM*, 43(12):45–48, Dec. 2000. ISSN 0001-0782. doi: 10.1145/355112.355122.

J. B. Schor and C. Fitzmaurice. Collaborating and connecting: the emergence of the sharing economy. *Handbook of Research on Sustainable Consumption*, pages 410–425, 2015. doi: https://doi.org/10.4337/9781783471270.00039.

B. Stvilia, M. Twidale, L. Gasser, and L. Smith. Information quality discussions in wikipedia. 01 2005.

M. ter Huurne, A. Ronteltap, R. Corten, and V. Buskens. Antecedents of trust in the sharing economy: A systematic review. *Journal of Consumer Behaviour*, 16(6):485–498, 2017. doi: 10.1002/cb.1667.

R. Thakur. Customer engagement and online reviews. *Journal of Retailing and Consumer Services*, 41(C):48–59, 2018.

A. Thierer, C. Koopman, A. Hobson, and C. Kuiper. How the internet, the sharing economy, and reputational feedback mechanisms solve the " lemons problem ". 05 2015. doi: 10.2139/ssrn.2610255.

S. Tobón, R.-A. R. José L, and J. Garcia-Madariaga. Gamification and online consumer decisions: Is the game over. *Decision Support Systems*, 128, 01 2020. doi: 10.1016/j.dss.2019.113167.

S. Utz, P. Kerkhof, and J. van den Bos. Consumers rule: How consumer reviews influence perceived trustworthiness of online stores. *Electronic Commerce Research and Applications*, 11(1):49 – 58, 2012. ISSN 1567-4223. doi: https://doi.org/10.1016/j.elerap.2011.07.010. Special Issue: CAT Tournament.

Y. D. Wang and H. H. Emurian. An overview of online trust: Concepts, elements, and implications. *Computers in Human Behavior*, 21(1):105 – 125, 2005. ISSN 0747-5632. doi: https://doi.org/10.1016/j.chb.2003.11.008.

S. Weinschenk. *Neuro Web Design: What Makes Them Click?* Voices That Matter. Pearson Education, 2009. ISBN 9780321648761. URL `https://books.google.nl/books?id=va7XOSAWZTMC`.

P. P. R. Zimmermann. *The official PGP user's guide*. Cambridge, Mass. : MIT Press, 1995. ISBN 0262740176 (pbk. : alk. paper). Includes bibliographical references and index.

Appendices

A   Coffee machine reviews

In the first experiment (as described in section 5.1.5), a variety of coffee machines with respective reviews and credibility features were presented to the participants. In this appendix, these reviews are shown. All credibility features that were tested are listed below:

1. Having a link to personal social media

2. Being a verified user (Verified as being authentic by a platform)

3. Being an endorsed user (E.g. an Airbnb 'superhost')

4. User feedback

5. Including a picture in the review

6. Having a personal avatar

1

Coffee machine Krups

€150

Marcus Bretagnio
June 2019
★★★★⯪

Wonderful how easy to use this thing proved to be. You can do other things while this brews your coffee for you. Grinds beans, packs the puck, does the brew, ejects the puck, and continues as directed. Used coffee grounds hopper is super easy to empty, while the drip tray is just as easy to rinse out when you are finished. LOVE THIS MACHINE!

2

Coffee machine Gaggia Anima

€150

Elizabeth Rouler ✔
June 2019
★★★★⯪

This machine exceeded all of my expectations. The lattes are delicious, even when using generic espresso. The frothing wand turned out to be the biggest surprise, as it heats the milk quickly without the loud noise of other machines that I've used.

3

€150

Coffee machine Philips EP2220

Deborah van Putten
July 2019
⭐⭐⭐⭐⯨

Endorsed

Had this machine a month now and am very happy it is part of our morning routine. If you take care of this machine as the instructions suggest, this machine will serve flawlessly for many years, I believe. The machine is very easy to clean per the instructions, I expect no trouble in the future. If anything goes wrong I'll come back and edit this review, but for today, this was a very worthwhile purchase.

4

€150

Coffee machine DeLonghi

Kathy Brown
July 2019
⭐⭐⭐⭐⯨

We spent more time looking at espresso machines than we did looking at our new house... Anyway, after much back and forth, we went with this one. It does not disappoint. It has been used daily (or more) for 1.5 years and still going strong. Great machine!

👍 Helpful Review | 12

5

€150

Coffee machine Philips HD8827

Joris Akkerman
June 2019
⭐⭐⭐⭐⯨

This is a great machine. It is easy to use and makes a high quality espresso. The frother is also very good. I have had it for 6 months and no issues. It is also fairly compact and doesn't take up much counter space. I included a picture of me making my coffee this morning!

6

€150

Coffee machine Caffe Corso

Gerda de Vries
July 2019
⭐⭐⭐⭐⯨

Insanely happy with this machine! Upon receiving I was a tad overwhelmed with the instruction manual and the thought of how often it needed to be cleaned/maintained but honestly it's been a breeze! The coffee that is made with this machine is superb compared to local cafes and I love the many options available; strength, brew time, size, etc. I've only used whole beans, so I can't speak to how well it is with pre-ground coffee, but I can't imagine it's much different! So happy with this!

B Results experiment 1

The results of the first experiment are posted in JSON format here. Some added notes: gender of the participants is either 1 (male) or 2 (female). In some answers there appears a "\n" in the string. This is the code for a linebreak.

```
1 {"gender":"2","age":"56","coffeemachine":"1","coffeemachinetext":"Best brand in
    my opinion","fullname":"2","nickname":"4","socialmedia":"4","avatar":"3","
    verified":"5","endorsed":"4","userfeedback":"3","sellerresponse":"3","
    picture":"2","privacy":"5","privacytext":"Do you mean if I was worried in
    this test? No\n\nDo you mean how worried I am in general while writing
    reviews? I keep my privacy as much as possible. ","prompt":"2","prompttext"
    :"If I think it is relevant to the owner or to other travelers. ","discount
    ":"1","discounttext":"Discount are always a hassle to reimbourse. ","
    credits":"2","creditstext":" I m  a super genius member with booking.com
    and like the conveniences. This is automatically done. ","doubleblind":"1",
    "doubleblindtext":"I trust it will be good. ","reflection":"1","
    reflectiontext":"I have a good imagination}
2 {"gender":"1","age":"22","coffeemachine":"5","coffeemachinetext":"the added
    image made me feel the review was more authentic. the time mentioned of
    using the device helped anchor that review in my mind. Even though it was '
    only' 6 months, the time seemed more believable than reviewing a product
    after over a year of ownership","fullname":"2","nickname":"1","socialmedia"
    :"3","avatar":"2","verified":"3","endorsed":"4","userfeedback":"3","
    sellerresponse":"3","picture":"5","privacy":"4","privacytext":"using real
    image and name is something I wouldn't want to do, and I don't feel like
    that's something I'd look at from other reviewers.\nThe concern for privacy
     holds even more strongly for linked social media accounts, as I feel the
    added accountability of the person reviewing doesn't make up for the
    privacy lost","prompt":"1","prompttext":"I don't find it worth the effort,
    especially since their review platform is skewed to make you leave a high
    star review, so I don't feel like my opinions can be sufficiently expressed
    . ","discount":"3","discounttext":"assuming the discount is sufficient, and
     I know I'll be using the platform again, I have a lot less to lose in this
     interaction","credits":"2","creditstext":"I'm not sure what the extra
    perks are, but without research, I doubt I'd expect them to be worth the
    effort ","doubleblind":"1","doubleblindtext":"I likely met their host, so I
     don't need to know more, and they are, like the guests, pushed towards
    positive reviews, so I rarely feel that they are genuine (outside of
    clearly negative experiences) ","reflection":"1","reflectiontext":"I had
    experience with some of the examples, and it was easy to extrapolate the
    rest from it"}
3 {"gender":"2","age":"22","coffeemachine":"1","coffeemachinetext":"Not entirely
    sure why. All reviews are super positive. I find it difficult to make a
    decision based solely on other p e o p l e s reviews, I rather look at the
```

differences in the machines and what I prefer when making a decision","
fullname":"2","nickname":"2","socialmedia":"4","avatar":"3","verified":"4",
"endorsed":"4","userfeedback":"3","sellerresponse":"3","picture":"4","
privacy":"5","privacytext":" I m not sure to what extent my own privacy
matters while looking at other p e o p l e s reviews","prompt":"2","prompttext
":"Usually not, only if I feel like I have something useful to add, then it
 might trigger me to share it. So only when I perceived something as rather
 negative that I really want to point out, or something extraordinary
positive that I want to address","discount":"3","discounttext":"Everyone
likes a discount, and leaving a review is a low effort way to get that
discount","credits":"3","creditstext":"I think so, if there is something
concrete that I I can get out of leaving a review, then it motivates me to
do it, since I get some sort of reward for it","doubleblind":"2","
doubleblindtext":"Ooh that depends on my curiosity levels haha! But I guess
 so. If I had a positive experience, I d think I get a positive review
myself, so then i t s nice to read something nice about yourself and
 i t s not much effort to leave a nice feedback for them. Although I m
not sure how I would act once I left a review myself and then find out my
experience and the experience of the host d o n t match, because that can
lead to negativities","reflection":"2","reflectiontext":"In the last one it
 was quite difficult to estimate my own behaviour, the one where the host
would leave a review as well. Because I really d o n t know how I would
react, I think it differs when I had a positive or negative experience, but
 still find it difficult to imagine"}

4 {"gender":"1","age":"26","coffeemachine":"5","coffeemachinetext":"Joris
Akkerman sound like the most trustworthy reviewer, because he has owned his
 machine for 6 months","fullname":"1","nickname":"3","socialmedia":"4","
avatar":"3","verified":"4","endorsed":"4","userfeedback":"3","
sellerresponse":"3","picture":"4","privacy":"4","privacytext":" ","prompt"
:"3","prompttext":"If I like the place, I feel like I can help the owners
and future travellers by letting them know the place is nice. It's nice to
share a positive experience ","discount":"3","discounttext":"I like
discounts","credits":"3","creditstext":"I like privileges","doubleblind":"3
","doubleblindtext":"They already got me with this one a few times","
reflection":"3","reflectiontext":" "}

5 {"gender":"1","age":"25","coffeemachine":"6","coffeemachinetext":"Gerda ziet er
 wel uit alsof ze betrouwbaar is. En als mensen geen pf hebben neem ik het
automatisch minder sirieus. joris maakt goeie 2e plaats.","fullname":"3","
nickname":"3","socialmedia":"4","avatar":"5","verified":"2","endorsed":"4",
"userfeedback":"2","sellerresponse":"2","picture":"2","privacy":"4","
privacytext":"ik geef zelf vrijwel nooit reviews. dus weinig","prompt":"1",
"prompttext":"spam. tenzij het een soort uitwisseling van een beoordeling
is. dan werkt het opzich wel weer.","discount":"3","discounttext":"money.",
"credits":"3","creditstext":"profit.","doubleblind":"3","doubleblindtext":"

curiosity. ","reflection":"1","reflectiontext":"ill leave a review if there
's something in it for me."}

6 {"gender":"1","age":"26","coffeemachine":"5","coffeemachinetext":"Real picture
of the product convinced me that the review is probably authentic and not
fake","fullname":"1","nickname":"1","socialmedia":"1","avatar":"3","
verified":"2","endorsed":"3","userfeedback":"2","sellerresponse":"1","
picture":"4","privacy":"4","privacytext":"I don't like having images of
myself or my name in public. Not everyone has to know that I bought
something","prompt":"1","prompttext":"you have to earn my review:\n- do
what is expected\n- do more than expected\n- have really terrible service",
"discount":"2","discounttext":"Depends on the discount... high discount
will be likely. But something like you will get 5 euro when you spend 100
isn't","credits":"3","creditstext":"Long term benefit","doubleblind":"2","
doubleblindtext":"on my mood","reflection":"1","reflectiontext":"clear
explanations"}

7 {"gender":"1","age":"23","coffeemachine":"3","coffeemachinetext":"I was
doubting between machine two and three, but the design of three looks
visually more pleasing to me (minimalistic, clean and not too edgy). What
sure helps is that some cups of coffee are displayed standing on the
machine.","fullname":"2","nickname":"2","socialmedia":"1","avatar":"1","
verified":"1","endorsed":"2","userfeedback":"4","sellerresponse":"5","
picture":"5","privacy":"1","privacytext":"I do not really get the question
to be honest","prompt":"3","prompttext":"Initially I would think NO, but
than I remember myself leaving a review to an AirBnb booking I really
enjoyed after they notified me to leave a review.","discount":"1","
discounttext":"I don't think so actually, because I usually do not plan my
next holiday so soon and especially because in general it annoys me if they
 urge you to do things like that. That specific reason gives me more reason
 to not do it. Furthermore the internet is so full of advertisements and
promotion actions, if I would really want to have a minor discount on my
booking I could probably find it else where, whenever I feel like planning
my next holiday.","credits":"1","creditstext":"Although (gamification)
theory states that those kind of reward systems stimulate people to perform
 actions, I think it is shit. I have been working on a project for three
months full-time, in which I had to develop a education game with a team.
We had to integrate that gamification theory as well, but every time I was
working with- or reading about it, it felt like it was not working.","
doubleblind":"1","doubleblindtext":"Holiday is over, I would be interested
in what he had to say about me if they wouldn't force me to do anything to
read it.","reflection":"1","reflectiontext":"Questions and situations were
clear and understandable. "}

8 {"gender":"1","age":"20","coffeemachine":"2","coffeemachinetext":"i like to
know how long a product survives and i trust a revieuw more if the product
is used longer, in this revieuw the the product is 1,5 years old and stil

going strong. and price isnt a factor in this situation","fullname":"1","
nickname":"1","socialmedia":"1","avatar":"3","verified":"3","endorsed":"3",
"userfeedback":"2","sellerresponse":"2","picture":"3","privacy":"4","
privacytext":"i find my privice fairly important, so i dont want to use my
real name on sites. for reviews i mostly on sites where is that is not as
import like reddit or tweakers. some one with a funny name can have a
revieuw just as good as some one that uses its real name","prompt":"1","
prompttext":"i dont like writing a revieuw because i'm logged in with my
main email. and i dont want this revieuw to be able to be linked back to me
","discount":"2","discounttext":"is stil dont like writing a revieuw but a
discount is nice. it depents if i am in a good mood and if i really liked
the airbnb","credits":"1","creditstext":"i dont like writing revieuws and
especially is my profiele is public. i dont care for badges and not having
al the priveleges from the start would only make me not use this platform.
","doubleblind":"1","doubleblindtext":"no i dont want to be pressured in
writing in a review and this feels like i would get a penalty in the form
of a bad revieuw by the host if i dont write a revieuw","reflection":"1","
reflectiontext":"no the situations where very clear and very plausible"}
9 {"gender":"1","age":"24","coffeemachine":"6","coffeemachinetext":"This was the
only machine with a photo of the person who gives the review. ","fullname":
"4","nickname":"1","socialmedia":"4","avatar":"5","verified":"5","endorsed"
:"5","userfeedback":"5","sellerresponse":"5","picture":"4","privacy":"2","
privacytext":"It did not pop into my mind about my privacy online when I
was answering the questions above.","prompt":"2","prompttext":"Not too sure
 why I would help Airbnb with my effort.","discount":"3","discounttext":"
Well, this gives me such a nice value to use airbnb again!","credits":"1","
creditstext":"I wouldn't know why I would get more priveleges if I give
reviews. Why don't I get all the privileges in the beginning of using?
Feels like they are hiding things from me.","doubleblind":"3","
doubleblindtext":"I am triggered to know what the other has said about me "
,"reflection":"1","reflectiontext":"It was easy to imagine every situation"
}
10 {"gender":"2","age":"22","coffeemachine":"2","coffeemachinetext":"Fancy design,
 italiaans merk (I guess) en aangezien daar goede koffie vandaan komt laat
ik me beinvloeden. ","fullname":"2","nickname":"1","socialmedia":"1","
avatar":"2","verified":"4","endorsed":"2","userfeedback":"1","picture":"1",
"prompt":"1","prompttext":"Te schreeuwerig, ik wil alleen een review
achterlaten als ik gebruik heb gemaakt van de site. Ze laten het zo
wanhopig klinken. ","discount":"3","discounttext":"KORTING ben hollander he
, gratis dingen. ","credits":"2","creditstext":"Het is wel fijn als
verhuurders in de toekomst mij als betrouwbaar beschouwen. ","doubleblind":
"1","doubleblindtext":"Super irritant, maar ook wel logisch want ik wil
zelf een eerlijke review schrijven en me niet laten be nvloeden door wat
hij/zij zegt. Aan de andere kant heb ik weleens gehad da tmijn review nooit

zichtbaar werd omdat de ander mij niet wilde reviewen. ","reflection":"1",
"reflectiontext":"Ik kan alles. "}

11 {"gender":"1","age":"16","coffeemachine":"3","coffeemachinetext":"it is
endorsed(whatever the fuck that means) and it is a brand I recognise. The
customer also said it has the ability to edit the review which is positive
in my eyes.","fullname":"4","nickname":"1","socialmedia":"4","avatar":"4","
verified":"5","endorsed":"4","userfeedback":"3","sellerresponse":"4","
picture":"4","privacy":"5","privacytext":"even though i am concerned or my
own privacy i do want to trust that real people review stuff. That is is
not review bombed or that there are fake reviews precent","prompt":"2","
prompttext":"if they review back I would. If they dont have enough reviews
that represent my feelings about my stay I would leave a review.","discount
":"3","discounttext":"Dutch free money","credits":"2","creditstext":"no and
 yes. I dont like this system it presures people into reviewing for better
stuff. More stays = more reviews = better places, i dont like a pay to win
principle. But you are kinda obliged to join in.","doubleblind":"3","
doubleblindtext":"curiousness ","reflection":"2","reflectiontext":"Not
realy"}

12 {"gender":"2","age":"28","coffeemachine":"2","coffeemachinetext":"Because I
like lattes and it mentions that they turn out good!","fullname":"4","
nickname":"3","socialmedia":"4","avatar":"3","verified":"4","endorsed":"5",
"userfeedback":"4","sellerresponse":"5","picture":"4","privacy":"5","
privacytext":"If I was the person writing the review then I would not
reveal any information about myself. I would try to be objective and
accurate and not sound fake. That's also how I judge a review ","prompt":"1
","prompttext":"I don't leave reviews unless I really wanna promote someone
 or something mostly because I'm lazy","discount":"3","discounttext":"If
there's benefit I'm less lazy","credits":"3","creditstext":"If there's
benefit I'm less lazy","doubleblind":"3","doubleblindtext":"Curiosity works
 really well on me","reflection":"1","reflectiontext":"Cuz it had happened
before"}

13 {"gender":"1","age":"25","coffeemachine":"2","coffeemachinetext":"Seems like a
real person because of the icon next to the name. Seems less like a fake
review.","fullname":"3","nickname":"4","socialmedia":"4","avatar":"3","
verified":"4","endorsed":"4","userfeedback":"1","sellerresponse":"1","
picture":"2","privacy":"4","privacytext":"Privacy immediately seemed
important with the first question. Seeing a real name helps a lot with the
credibility of a review, but personally I would not want to post a review
under my real full name.","prompt":"2","prompttext":"Depends on if I feel
like it, I am more inclined to review is the experience was bad than when
it was good.","discount":"3","discounttext":"Because I like discounts! That
 would give me some motivation, if the review process is too cumbersome
though I would still not follow through.","credits":"1","creditstext":"I
 d o n t care a lot about online reputation.","doubleblind":"2","

doubleblindtext":"If I had a connection with the host I might be interested in what message he/she left for me.","reflection":"1","reflectiontext":"Because I have experienced many of the situations in one form or another."}

14 {"gender":"1","age":"25","coffeemachine":"5","coffeemachinetext":"Joris seems like a down to earth guy who probably speaks the truth. 6 months and no issues klinkt ook vrij goed en meten is weten. Gerda zit in de overgang en vindt alles mooi en is dus semi betrouwbaar.","fullname":"5","nickname":"1","socialmedia":"4","avatar":"5","verified":"3","endorsed":"4","userfeedback":"3","sellerresponse":"1","picture":"5","privacy":"5","privacytext":"besides the fact that i am a top-notch influencer, did my concern for my own privacy not influenced my answers given above. I do not write reviews myself, but if i would want to write a review to help other people out i wouldn't mind showing my face and name. If i would be completely anonymous i wouldnt take myself serious to be honest. But that could also be a reason why i don't write reviews. I don't like to put my face online really. ","prompt":"2","prompttext":"writing reviews is not a great hobby of mine. so i would not be very triggered to write one. But. if the host was really nice, humble etc. i would feel bad not leaving a review and this could remind me to leave one. so yea it could trigger me but only if the experience was good. ","discount":"3","discounttext":"you had my curiosity but now you have my attention. money talks.","credits":"1","creditstext":"Too much of a hassle","doubleblind":"3","doubleblindtext":"What did that mfker say about me","reflection":"1","reflectiontext":"I'm not sure if i understand this question completely. But i think it wasn't hard to imagine the situations because i've been in similar situations before.\nfeeling bad about not leaving a review or not wanting to do stuff for a badge online, or stuff like buy for 100 euro at coffeeshop X and get 10 cents of when you have a volle spaarkaart. takes too long and i would probably forget i started it."}

15 {"gender":"1","age":"55","coffeemachine":"4","coffeemachinetext":"Gives more confidence than the other reviews","fullname":"4","nickname":"1","socialmedia":"3","avatar":"3","verified":"3","endorsed":"3","userfeedback":"3","sellerresponse":"4","picture":"3","privacy":"4","privacytext":"Not, confident that you will use this properly ","prompt":"2","prompttext":"Not in case of a moderate stay","discount":"3","discounttext":"A postive reward .... ","credits":"3","creditstext":"Positive reward... ","doubleblind":"3","doubleblindtext":"Intriguing., makes curious. ","reflection":"1","reflectiontext":"Happend several times only weeks ago"}

16 {"gender":"1","age":"66","coffeemachine":"4","coffeemachinetext":"na 1,5 jaar nog steeds enthousiast","fullname":"4","nickname":"1","socialmedia":"3","avatar":"4","verified":"5","endorsed":"4","userfeedback":"4","sellerresponse":"4","picture":"3","privacy":"5","privacytext":"privacy is belangrijk, bijv. niet lastig gevallen worden door andere aanbieders","prompt":"3","prompttext":"om volgende gebruikers van dienst te zijn, op

dezelfde manier als ik reviews van anderen waardeer","discount":"3"," discounttext":"om volgende gebruikers van dienst te zijn, op dezelfde manier als ik reviews van anderen waardeer, de beloning stimuleert dat"," credits":"3","creditstext":"om volgende gebruikers van dienst te zijn, op dezelfde manier als ik reviews van anderen waardeer, de beloning stimuleert dat","doubleblind":"2","doubleblindtext":"is minder stimulerend"," reflection":"1","reflectiontext":"-"}

17 {"gender":"2","age":"22","coffeemachine":"1","coffeemachinetext":"It looks nice and stylish","fullname":"4","nickname":"3","socialmedia":"2","avatar":"4", "verified":"4","endorsed":"4","userfeedback":"4","sellerresponse":"3"," picture":"3","privacy":"4","privacytext":"I did not really think about is. I think all above is important but I wont do it myself","prompt":"2"," prompttext":"Depends on how good or bad it was. I mostly don't leave a review if it is moderate","discount":"3","discounttext":"Money money money" ,"credits":"3","creditstext":"It helps if you get something back for your effort","doubleblind":"3","doubleblindtext":"Just really curious"," reflection":"1","reflectiontext":"I have hosted on air bnb so that makes it easier"}

18 {"gender":"2","age":"22","coffeemachine":"4","coffeemachinetext":"durability"," fullname":"2","nickname":"4","socialmedia":"1","avatar":"2","verified":"2", "endorsed":"1","userfeedback":"4","sellerresponse":"5","picture":"2"," privacy":"4","privacytext":"I wondered why I had to send my gender mostly. I would not like personally to have my name somewhere, I'd rather pick a nickname. Also photos can give much info, so I would not send them with a review myself.","prompt":"2","prompttext":"If I went and the stay was about the same as the reviews I would not, otherwise I would and I would be more inclined with such a sign","discount":"1","discounttext":"I would not do that, since that feels like rigging your reviews","credits":"2"," creditstext":"If i would care about the platform maybe","doubleblind":"1"," doubleblindtext":"Feels like I'm making a test, would not like that"," reflection":"1","reflectiontext":"Very clear"}

19 {"gender":"2","age":"37","coffeemachine":"5","coffeemachinetext":"An extra picture says more than a 1000 words to me.","fullname":"4","nickname":"1"," socialmedia":"2","avatar":"4","verified":"4","endorsed":"4","userfeedback": "4","sellerresponse":"3","picture":"4","privacy":"4","privacytext":"It didn 't. I don't believe there is much privacy when going online. So, don't post anything you don't want others to see, read or find. ","prompt":"2"," prompttext":"I only write reviews if something is really great or really bad, because I don't think standard reviews like \"it was okay\" help people.","discount":"3","discounttext":"Because humans are opportunistic and so most of us are motivated by personality gain.","credits":"3"," creditstext":"Again, there is something to gain and that motivates people. ","doubleblind":"3","doubleblindtext":"I would be to curious about what the other person had to say about me.","reflection":"1","reflectiontext":"No,

not at all. I already thought about things like this myself, because I am
very concious about what I post online."}

20 {"gender":"1","age":"25","coffeemachine":"4","coffeemachinetext":"The positive
review is endorsed by multiple people. ","fullname":"4","nickname":"1","
socialmedia":"3","avatar":"5","verified":"4","endorsed":"5","userfeedback":
"5","sellerresponse":"2","picture":"3","privacy":"4","privacytext":"Seeing
a last name, makes a profile more credible. But I would be a little
hesitant to share my last name on any given platform","prompt":"2","
prompttext":"Depends on the service or product I got, and how content I am
with it. Another factor might be the expected effort for leaving a review.
If I see that it simply is a star rating or providing a single grade, I
would do it. Writing a review is time-consuming, and I might skip it if I
don't have a point to make or something to gain from it. ","discount":"3","
discounttext":"Airbnb is a service I use regularly (couple times a year),
the discount is, therefore, likely to be used which makes leaving a review
worth it for me.","credits":"3","creditstext":"This incentive is worth it,
and leaving a review is a consideration especially if it is low effort.
Although the benefit and value of an 'improved profile' is less obvious
compared to a direct discount.","doubleblind":"3","doubleblindtext":"
Curiosity. Certainly, if I knew the review was visible to other hosts of
potential future stays.","reflection":"1","reflectiontext":"I am familiar
with Airbnb and have come across a lot of 'leave a review' messages. I
reflected on at which times I left a review, and when I didn't and why.\n\n
"}

21 {"gender":"1","age":"25","coffeemachine":"5","coffeemachinetext":"It a clear
and compact review\n\nNo fan of the fb logo at the first option\n","
fullname":"2","nickname":"3","socialmedia":"1","avatar":"3","verified":"4",
"endorsed":"4","userfeedback":"4","sellerresponse":"2","picture":"3","
privacy":"5","privacytext":"Not so much","prompt":"1","prompttext":"Sounds
desperate ","discount":"3","discounttext":"Not a lot of work for a discount
","credits":"2","creditstext":"Not a big user of airbnb, so saving up
badges wouldn't go fast","doubleblind":"1","doubleblindtext":"Not
interested ","reflection":"3","reflectiontext":"I would leave a review in
case of an extraordinary experience, if very bad to warn others and if very
 good to reward the owners/supplier. Otherwise not very interesting to
leave a review. "}

22 {"gender":"1","age":"58","coffeemachine":"6","coffeemachinetext":"more personal
 due to the picture and the enthousiastic test. It alsohelps that two
filled coffee cups are present.","fullname":"4","nickname":"1","socialmedia
":"3","avatar":"4","verified":"5","endorsed":"4","userfeedback":"2","
sellerresponse":"3","picture":"1","privacy":"4","privacytext":"it did not.
Was temporarily not aware of my privacy","prompt":"1","prompttext":"too
many of such reviews","discount":"1","discounttext":"no as this is just a
waytolure you into another buy","credits":"1","creditstext":"same answer

as above","doubleblind":"1","doubleblindtext":"why would I like to know
  what a host thinks about me or my behaviour? I am just renting a place.","
  reflection":"1","reflectiontext":"no, I often get this kind of requests"}
23 {"gender":"2","age":"55","coffeemachine":"2","coffeemachinetext":"Love a good
  latte!!","fullname":"3","nickname":"3","socialmedia":"4","avatar":"3","
  verified":"4","endorsed":"4","userfeedback":"5","sellerresponse":"5","
  picture":"3","privacy":"4","privacytext":"Not so much of a concern","prompt
  ":"3","prompttext":"It is nice to leave a review if the stay was okay, to
  help people out to get new gasts","discount":"3","discounttext":"Always
  nice to get a discount","credits":"3","creditstext":"Iike to stay in
  airbnbs, so when you get more special offers as in special overstay-places
  i would love that","doubleblind":"3","doubleblindtext":"Just curious","
  reflection":"1","reflectiontext":"I just imagine easily"}
24 {"gender":"2","age":"23","coffeemachine":"6","coffeemachinetext":"I hesitated
  between nr 4 and 6: They seem to be very honest reviews. But 4 doesn't say
  anything about cleaning: I am interested in a machine that is easy to clean
  . Though it interested me that the other machine was already in use for 1.5
   years, the very personal and extensive review of machine 6 convinced me
  more.","fullname":"2","nickname":"1","socialmedia":"2","avatar":"3","
  verified":"3","endorsed":"3","userfeedback":"1","sellerresponse":"1","
  picture":"1","privacy":"4","privacytext":"I understand that seeing \"the
  real person\" in terms of an image or name would make something more
  credible, but it is not of real importance to me because I assume that
  people do not want their info to be shared like that. I would even suggest
  that more credible people would care more about their privacy, but that is
  an ungrounded assumption :) ","prompt":"1","prompttext":"This is basically
  because of my notification experience of Airbnb. I mostly feel a bit
  spammed by them which makes me ignore their notifications :)","discount":"2
  ","discounttext":"Hmmm still not so willing to help to be honest, but if I
  would be planning a trip using airbnb I would absolutely consider.. ","
  credits":"2","creditstext":"In this case it wouldn't be so beneficial to me
   since I am not considering renting my own space to others (and I think
  that's where badges and privileges would come in most handy). And I am
  happy with the platform as it is right now without the extra privileges..
  But then again if I am planning a trip or feel bored anyway this could be a
   reason for me to help them out.","doubleblind":"3","doubleblindtext":"Yeah
   that's the kind of motivation that triggers me! Whoops :) ","reflection":"
  1","reflectiontext":"I have experience using airbnb so I could just imagine
   the situation as I've experienced it (so not even imagine it but just
  remember it)"}
25 {"gender":"2","age":"53","coffeemachine":"4","coffeemachinetext":"Former good
  experience with this brand.\nLooks good","fullname":"2","nickname":"2","
  socialmedia":"2","avatar":"2","verified":"4","endorsed":"4","userfeedback":
  "2","sellerresponse":"3","picture":"4","privacy":"4","privacytext":"I dont

give much value to reviews on websites. I don't ever post a review of a
product on a website because of privacy concerns. I do find pictures
helpfull sometimes, but i don't post them","prompt":"1","prompttext":"I don
't like to be bothered with this things.","discount":"1","discounttext":"
Still i dont like to review and appear on line","credits":"1","creditstext"
:"Still not amused","doubleblind":"1","doubleblindtext":"Less amused, i
dont like to be revieuwd and have no interest in what a host saus about me!
","reflection":"1","reflectiontext":"Not hard, it happens all the time that
 im asked to revieuw a product or place. I just dont like to be botherd
with it, no matyer what they offer."}

26 {"gender":"2","age":"51","coffeemachine":"4","coffeemachinetext":"Reviews are
positive, machine looks fine.","fullname":"3","nickname":"1","socialmedia":
"1","avatar":"3","verified":"3","endorsed":"2","userfeedback":"2","
sellerresponse":"1","picture":"2","privacy":"5","privacytext":"Not, not at
this point. Sometimes I do not like mentioning my name in a review, because
 of the privacy.","prompt":"3","prompttext":"Next visitors will also read
reviews before they decide, so I will leave a review to help other people
to make a decision.","discount":"2","discounttext":"It would not be an
extra reason for me to leave a review. Maybe feeling a little bit
    pushed   to do so.","credits":"1","creditstext":" D o n t like this. ","
doubleblind":"1","doubleblindtext":"Feeling of being pushed to leave a
review. ","reflection":"1","reflectiontext":"On Social media e.g. you are
experiencing this way of reacting. "}

27 {"gender":"2","age":"22","coffeemachine":"3","coffeemachinetext":"They will
edit the review if i t s not accurate anymore","fullname":"2","nickname":"
3","socialmedia":"1","avatar":"1","verified":"4","endorsed":"4","
userfeedback":"4","sellerresponse":"3","picture":"1","privacy":"5","
privacytext":"No need to have a picture or full name. If someone has a
weird nickname I take it less seriously ","prompt":"2","prompttext":" I m
more likely to leave a review if i t s a bad one.","discount":"1","
discounttext":"Because I d o n t know yet if I will book there again and by
 that time I have probably forgotten about it. ","credits":"3","creditstext
":"More options is better. ","doubleblind":"3","doubleblindtext":"Because I
 would be curious what the host has written about me. ","reflection":"1","
reflectiontext":"-"}

28 {"gender":"2","age":"24","coffeemachine":"4","coffeemachinetext":"The user
explains that she has used the machine for 1,5 years, which is longer than
the other users indicate to have used their machines without problems","
fullname":"1","nickname":"4","socialmedia":"1","avatar":"3","verified":"4",
"endorsed":"2","userfeedback":"1","sellerresponse":"3","picture":"5","
privacy":"5","privacytext":"This is one of the reasons I d i d n t indicate
that the full name of socials are important indeed.","prompt":"3","
prompttext":"I know reviews are important for services provided online, so
if I have the time, I would review","discount":"3","discounttext":"I think

I will be using Airbnb again so this would be helpful for myself as well","
credits":"1","creditstext":"I d o n t really care about having more
privileges than just being able to use the platform in its free form","
doubleblind":"3","doubleblindtext":"I would be curious","reflection":"1","
reflectiontext":"I think this is something that happens In real life often
enough to be able to imagine it"}
29 {"gender":"1","age":"55","coffeemachine":"3","coffeemachinetext":"Ziet er
simpel uit en is van een bekend merk","fullname":"4","nickname":"3","
socialmedia":"4","avatar":"4","verified":"4","endorsed":"4","userfeedback":
"4","sellerresponse":"4","picture":"4","privacy":"3","privacytext":"Privacy
is belangrijk maar is een afweging in het delen en beschikbaar hebben van
informatie","prompt":"2","prompttext":"Waarschijnlijk wel, het ligt eraan
hoe eenvoudig het is","discount":"2","discounttext":"Ook hier geldt dat
eenvoud belangrijk is","credits":"2","creditstext":"De beloning is niet de
belangrijkste drijfveer, het gemak en de serieusheid van de website is
belangrijker","doubleblind":"2","doubleblindtext":"Hangt van de gastheer af
","reflection":"2","reflectiontext":"Omdat er redelijk vaak om wordt
gevraagd oor allerlei websites en bijvoorbeeld ook Googel wanneer je
openbaar vervoer via maps gebruikt"}

C   Results experiment 2

The results of the second experiment are posted in JSON format here. Some added notes:
gender of the participants is either 1 (male) or 2 (female). In some answers there appears a
"\n" in the string. This is the code for a linebreak.

1 {"gender":"1","age":"27","platformuser":"1","platformusertext":"","
headphonesreviewcontent":" I got my fresh pair of Sony headphones last week
. The noise cancelling feature is great, living next to a cafe makes me
appreciate the clean music even more! I worked in audiotech for a few years
, and the guys that are working on this clearly know what they are doing. I
 would buy this again if it ever breaks.","headphonesreviewtext":"He seems
to be an expert","interestreviewtext":"Expert opinions","
middleexperiencetext":"Only if I got something more out of it","
simpleratingtext":"4 star rating. the middle star in a 5 star rating doesn'
t tell anything. it has the same effect as just not voting at all.","
fairratingtext":"I would not. I would only comment if I feel my opinion is
of added value. ","gamificationrewardtext":"If it brings extra benefits,
like being on top of a search query, sure!","gamificationchallengetext":"
Yes, but only if you were doubting to leave a review already. It would make
 the decision easier and also motivate to do your best. ","
gamificationmeaningfultext":"This is too far away, and often I am a bit
skeptical about these promises. So no. ","gamificationinteractivitytext":"

Only if it is a platform that I am really interested in and I am willing to invest time it as if it were a hobby. "}

2 {"gender":"2","age":"37","platformuser":"1","platformusertext":"","headphonesreviewcontent":{"gender":"F","name":"Marian Ketelburgh","picture":true,"review":" Last year I purchased them, and I have used it over the year. I love how balanced the sound of the headphones is, a very natural sound that fits all music. I have bought other devices from the same brand, they never dissapoint. They do what they must do, which is good. You can see what they look like in the picture that I added!"},"headphonesreviewtext":"De laatste die een goeie recensie deelde ","interestreviewtext":"Meer positieve reactie zet ik tegen over de negatieve en haal daar mijn conclusie uit, meer mijn gevoel op dat moment. ","middleexperiencetext":"Ja \n","simpleratingtext":"Ja ik vind de sterren erg makkelijk ","fairratingtext":"Nee ","gamificationrewardtext":"Nee dat vind ik raar","gamificationchallengetext":"Nee dat vind ik ook niet goed ","gamificationmeaningfultext":"Nee dat kan niet, het zou wel goed zijn maar ik vraag me af of mensen daar geen gebruik van gaan maken. ","gamificationinteractivitytext":"Nee nee nee"}

3 {"gender":"2","age":"29","platformuser":"1","platformusertext":"","headphonesreviewcontent":{"gender":"F","name":"Eva van Kampen","picture":false,"review":" I've had this product for about six months. The headphones are very durable, the material feels strong. When I dropped them, there wasn't a single scratch! I have bought other devices from the same brand, they never dissapoint. They do what they must do, which is good."},"headphonesreviewtext":"6 month seem to be a good time to test headphones, it seemed solid. Also full stars usually seem a little suspicious. Also user knows the brand.","interestreviewtext":"Time frame (not too early nor too long ago) \nStars (I like it if people are a little critical) \nHonest (I hate if it people give an overly enthusiastic opinion that seems fake)","middleexperiencetext":"Yes! Because I value an honest opinion, and I hope others do too. For me the middle sections is the most interesting. 5 stars or 1 star ratings are   overrated   \nThe   juicy   bits are in the middle. ","simpleratingtext":"I find just a like or dislike button to simple, life a i n t simple nor is a review. I like the star rating tbh. Makes it easier to filter out the bullshit. ","fairratingtext":"No but I would like to upvote that rating or give it a   thumbs   u p to show that I agree with it ","gamificationrewardtext":"Nope that is just another way of farming karma. I would question, if people actually do this because they want to or just to show off how important their opinion is. And why would that opinion be better than from someone without that badge. ","gamificationchallengetext":"Meh same reason as before, I would question their motive of writing a review (also if they get something for free how honest is that opinion) ","gamificationmeaningfultext":"Haha got me there, we need more trees!!! Nevertheless still would thing it would not be super

honest. Why do people always need a reward in order to do something? Even
if it is for the environment! ","gamificationinteractivitytext":"Not
necessarily. "}
4 {"gender":"2","age":"56","platformuser":"2","platformusertext":"Aibnb,
marktplaats","headphonesreviewcontent":{"gender":"M","name":"Bart
Ochtendschalker","picture":false,"review":" I've had this product for about
six months. The rich deep bass works really well for upbeat tracks, like I
'm standing right in the club. After looking at other options online, this
one seemed to be the best value for money. They do what they must do, which
is good."},"headphonesreviewtext":"He used the product for 6 months that's
not too long but long enough to know how the quality is","
interestreviewtext":"If people are happy with the product if it works the
way it is advertised ","middleexperiencetext":"Sometimes I do but often I
think it takes to much time or I haven't used the product long enough to
give a good review ","text":"I would do it easier I quess","fairratingtext"
:"Probably not only if I like the product very much or if the product is no
good I will tell what is good or not on the product","
gamificationrewardtext":"No I don't care much about these signs that other
people can see how good or popular you are","gamificationchallengetext":"
Maybe I will its a challenge to write an honest review even if you know
your change of getting to win is minimal","gamificationmeaningfultext":"
Absolutely and I will tell others to do so too","
gamificationinteractivitytext":"No if it gets personal I wouldn't give
reviews I don't like a discussion over products"}
5 {"gender":"2","age":"52","platformuser":"1","platformusertext":"","
headphonesreviewcontent":{"gender":"F","name":"Eva van Kampen","picture":
true,"review":" I bought my pair 1.5 years ago. The rich deep bass works
really well for upbeat tracks, like I'm standing right in the club. I
worked in audiotech for a few years, and the guys that are working on this
clearly know what they are doing. I would buy this again if it ever breaks.
You can see what they look like in the picture that I added!"},"
headphonesreviewtext":"After 1,5 yesr still happy with his buy","
interestreviewtext":"Long lasting products","middleexperiencetext":"It is
what it is sometimes t is middle experience.","text":"It is easier but also
not always 100% through.\n\nSometimes to.easy to do \n","fairratingtext":
"No","gamificationrewardtext":"No not important a reward or not ","
gamificationchallengetext":"No i put no energy in this kind of things","
gamificationmeaningfultext":"It should be but i.dont think i would do it","
gamificationinteractivitytext":"No to.much to do"}
6 {"gender":"1","age":"22","platformuser":"2","platformusertext":"social media
are seldomly used\ni use airbnb when travelling (2-3x per year)\n\n","
headphonesreviewcontent":{"gender":"M","name":"Geert Snoeibrug","picture":
true,"review":" One month ago I decided to get new headphones. The
headphones are very durable, the material feels strong. When I dropped them

, there wasn't a single scratch! I worked in audiotech for a few years, and the guys that are working on this clearly know what they are doing. I enjoy listening to my favorite songs with them! You can see what they look like in the picture that I added!"},"headphonesreviewtext":"the review was positive, it provided an image, and seemed to come from a reliable source", "interestreviewtext":"an image, as proof that the review is genuine.\nAdditionally, the language used by the user in the review can indicate how into that product category they are, which would affect how accurately I take their word","middleexperiencetext":"I almost never leave reviews, so I still wouldn't.\nI only make an exception for something that I would have a strong opinion on that I could elaborate on. ","simpleratingtext":"if it's a popup that I can get rid of with 1 click (whether it's the 'x', star rating, or lile/dislike) I would tend to give feedback if my experience had been positive so far","fairratingtext":"I'd rather just indicate my agreement with the review, the way it is done on Amazon (* people found this helpful) ","gamificationrewardtext":"if this badge will equal be being more in demand, and as a result earn more money (like on airbnb) then its worth it.\nbeing a well rated Amazon reviewer would not give you those benefits","gamificationchallengetext":"no, because that likely means the manufacturer determines the winner, so they pick good reviews, and then it just sounds like they're fishing for positive feedback"," gamificationmeaningfultext":"it's difficult to determine what well-written means, and it should have easy to fulfill criteria (like the little checkmarks when you're making a password, and the website indicates which requirements you've fulfilled).\nif I can know if my review is good enough prior to posting, I'd do it. otherwise, I'd worry my effort would be wasted ","gamificationinteractivitytext":"this would discourage me from leaving negative reviews, as they would be traced back to me. additionally, this would prevent me from reviewing products I don't want other people to know I've purchased"}

7 {"gender":"2","age":"55","platformuser":"1","platformusertext":""," headphonesreviewcontent":{"gender":"M","name":"Geert Snoeibrug","picture": false,"review":" I bought my pair 1.5 years ago. The noise cancelling feature is great, living next to a cafe makes me appreciate the clean music even more! I have bought other devices from the same brand, they never dissapoint. This is probably my new favorite pair of headphones!"}," headphonesreviewtext":"Long time experience","interestreviewtext":"Does the equipment do what iT should do","middleexperiencetext":"No I do not give my opinion.\nWhat you experience is personal","simpleratingtext":"Maybe, SER above","fairratingtext":"No","gamificationrewardtext":"No, takes me to much time","gamificationchallengetext":"No, see above"," gamificationmeaningfultext":"That Will help !"," gamificationinteractivitytext":"No"}

8 {"gender":"1","age":"55","platformuser":"2","platformusertext":"AirBnB,

Marktplaats?, ","headphonesreviewcontent":{"gender":"M","name":"Geert
Snoeibrug","picture":false,"review":" I bought my pair 1.5 years ago. The
headphones are very durable, the material feels strong. When I dropped them
, there wasn't a single scratch! This is the first time that I buy
headphones from this brand, but I am sure it won't be the last. They do
what they must do, which is good."},"headphonesreviewtext":"A comparisson
with other brands is made, longer experience is shared, material quality is
 mentioned","interestreviewtext":"longer term experience, quality compared
to others","middleexperiencetext":"Yes, to complete the picture for other
people","simpleratingtext":"the emotion-button would work fine for sevices"
,"fairratingtext":"I would have less urge.","gamificationrewardtext":"Not
for me, no added value..","gamificationchallengetext":"Difficult. What is
best? Who does judge?","gamificationmeaningfultext":"That would help to
stimulate!","gamificationinteractivitytext":"Will not work for me."}
9 {"gender":"1","age":"29","platformuser":"2","platformusertext":"Airbnb: ~ 2
times per year\nAmazon/Ebay/Marktplaats: ~ 5 times per year\n","
headphonesreviewcontent":{"gender":"F","name":"Eva van Kampen","picture":
false,"review":" I got my fresh pair of Sony headphones last week. I love
how balanced the sound of the headphones is, a very natural sound that fits
 all music. I have bought other devices from the same brand, they never
dissapoint. This is probably my new favorite pair of headphones!"},"
headphonesreviewtext":"The overall reviews (1-6) are positive, so I was
looking at the review that convinced me to buy the headphones rather than
the one convincing me not to. Review 6 provides information about the sound
 quality of the product, which to me is more important than e.g. durability
 or design. Moreover, Eva mentions other products of the manufacturer and
tells us to have an overall great experience with them.","
interestreviewtext":"Best bang for the buck generally, but who isn't? I
make my decisions based on brand name (with which I had good experience
before), price level, and the most important features that I use, i.e. I
don't buy products in a higher price level just for the sake of having more
 features than I'd likely need. ","middleexperiencetext":"I know how
important it is and yet I still barely provide reviews (if any). Most
likely because it takes a lot of time, or because I forget to do so. I'd
probably only rate if I had an either good or bad experience to inform
other potential buyers on the pros/cons that may not have been mentioned
yet. Generally, I only write positive reviews to support people I
personally met, such as tour guides or Airbnb hosts. But again, only if I
had a very good experience. ","simpleratingtext":"It would. Less time ->
larger chance of leaving a review.","fairratingtext":"I would not, because
it doesn't change the overall rating to which other people may draw
conclusions. ","gamificationrewardtext":"Although I would not be triggered
more than I am now, rewarding customers in general will definitely improve
platform engagement. It stimulates activity, which is always useful whether

  or not customers share positive experiences. ","gamificationchallengetext"
:"Good idea, but it is prone to subjectivity. Also, it will probably
stimulate people to only write very positive reviews. Or would you select a
 prize-winning review that burns your product to the ground? ;) ","
gamificationmeaningfultext":"Fantastic idea, but not feasible for most
companies. Will only work for the very large ones. I'd be excited if there
somehow is a way to realise it though.","gamificationinteractivitytext":"
This works if people are repetitive customers of a single company, i.e. are
 brand loyal. In that case your suggestion would be a great asset to the
customer intimacy and creates an invaluable platform of networking people."
}

10 {"gender":"1","age":"23","platformuser":"3","platformusertext":"Airbnb, Steam,
Aliexpress, Marktplaats Thuisbezorgd. I use them ones a week","
headphonesreviewcontent":{"gender":"M","name":"Geert Snoeibrug","picture":
false,"review":" I bought my pair 1.5 years ago. The headphones are very
durable, the material feels strong. When I dropped them, there wasn't a
single scratch! I bought something from this brand before, but was not
particularly happy with it. I still wanted to give them a chance with these
 headphones, though. I would recommend this product."},"
headphonesreviewtext":"I would buy it because customer 4 owns the product
for quite a while now and is clearly satisfied with it's durability. Also
he recommended to buy it.","interestreviewtext":"I search for different
reasons. First to find complains that are repeatedly in the comments. That
would be a good reason to search for somthing another product. Secondly I
would try to find reviews that compare the product with its competitors.
Just to check it wouldn't be a bad buy.\nAnd last but not least I would try
 to find reviews of people using the product for a long time. Just to be
sure it is durable.","middleexperiencetext":"Yeah I would. But most of the
time i don't realise how important that can be to others.","
simpleratingtext":"Personally I would like to have the 5 star rating. It
isn't hard to fill in but it is more accurate than thumbs up or down","
fairratingtext":"If i feel like leaving a review it wouldn't stop me if
someone else all ready left the same one.","gamificationrewardtext":"This
wouldn't motivatie me.","gamificationchallengetext":"I feel this would make
 the rating appear more positive than it should be. So no not for me.","
gamificationmeaningfultext":"Yeah I like this idea!","
gamificationinteractivitytext":"I think it would be good to be able to rate
 other reviews but the other stuff i would not care about that."}

11 {"gender":"1","age":"22","platformuser":"2","platformusertext":"Airbnb and Uber
 when on holiday. I also like to browse through United Wardrobe and
Marktplaats every now and then but have never bought anything.","
headphonesreviewcontent":{"gender":"M","name":"Bart Ochtendschalker","
picture":true,"review":" I got my fresh pair of Sony headphones last week.
I love how balanced the sound of the headphones is, a very natural sound

that fits all music. I worked in audiotech for a few years, and the guys that are working on this clearly know what they are doing. This is probably my new favorite pair of headphones! You can see what they look like in the picture that I added!"},"headphonesreviewtext":"This guy apparently knows his stuff about audio","interestreviewtext":"Someone who sounds reasonable and objective","middleexperiencetext":"I know the feedback is useful for other people but I never leave a review, because I'm too lazy. I feel like I have to write something when leaving a review since I hate it when people review a product with a star rating without the reasons for their rating. ","simpleratingtext":"My ideal rating system would look like being able to rate a few important aspects of the product with a star rating or something similar. So you don't have to write text and the star rating has more meaning than an overall star rating. ","fairratingtext":"I probably still would write a review (if writing reviews was something I did). But I wouldn't really think about the point you noticed in your first experiment.","gamificationrewardtext":"Yes, I think these rewards can be very helpful. It gives you an indication of the trustworthiness of someone.","gamificationchallengetext":"No, I think people would write reviews less honest. On the other hand, it does stimulate more reviews. ","gamificationmeaningfultext":"Yes, it's a win-win for you and the person who is reviewed. You do something meaningful for something you think is important and the owner of the product receives useful feedback.","gamificationinteractivitytext":"Not necessarily, I like to be anonymous when leaving reviews and such and don't see the added value of having a profile. I do think a feature where you can rate someones review or where you can indicate whether you agree or disagree with their review would add a lot of value. "}

12 {"gender":"2","age":"64","platformuser":"3","platformusertext":"Marktplaats\nAirbnb\nTripadvisor\nBooking.com\nKayak.com\n\nUsw.","headphonesreviewcontent":{"gender":"M","name":"Geert Snoeibrug","picture":false,"review":" I bought my pair 1.5 years ago. The noise cancelling feature is great, living next to a cafe makes me appreciate the clean music even more! I have bought other devices from the same brand, they never dissapoint. I would recommend this product."},"headphonesreviewtext":"1,5 jaar in gebruik, dus goede kwaliteit \nDe reden waarvoor je koptelefoons aanschaft is in orde: goed geluid\nEerdere ervaringen met het bedrijf geven vertrouwen\n","interestreviewtext":"Duurzaamheid van het product\nDoet het waarvoor je het wilt gebruiken","middleexperiencetext":"Ik denk dat ik die moeite niet zou doen.","simpleratingtext":"Voor mij is een combi het meest ideaal: sterren n toelichting. De eerste blik valt op het aantal sterren en vervolgens lees ik bij de reviewsja, dat zou voor mij niet uitmaken.","fairratingtext":"Ja, dat zou voor mij niet uitmaken.","gamificationrewardtext":"Ja, het geeft vertrouwen in het platform. Alsof ze er serieuzer mee omgaan dan wanneer dat niet gebeurt (waarschijnlijk een

marketing truc, maar ja, daar stink ik dan dus in).","
gamificationchallengetext":"Nee, daar kijk ik dan w l weer doorheen. Daar
doe ik niet aan mee.","gamificationmeaningfultext":"Nee, ook dit valt bij
mij, net als de vorige, onder het kopje   omkoperij   en wekt eerder
irritatie dan motivatie.","gamificationinteractivitytext":"Nee, het moet
niet teveel tijd/moeite kosten."}
13 {"gender":"2","age":"22","platformuser":"3","platformusertext":"Marktplaats 4x
per jaar\nThuisbezorgd 2x per maand\nAirbnb 1x per jaar\n","
headphonesreviewcontent":{"gender":"F","name":"Eva van Kampen","picture":
true,"review":" I got my fresh pair of Sony headphones last week. It feels
very comfortable on my head, I can wear it all day long. I bought something
 from this brand before, but was not particularly happy with it. I still
wanted to give them a chance with these headphones, though. I enjoy
listening to my favorite songs with them! I added a picture to my review to
 show what mine look like."},"headphonesreviewtext":"Omdat ze eerder al van
 hetzelfde merk had gekocht dus kan vergelijken.","interestreviewtext":"
Kort maar krachtig, pakkend.\n\nIk zoek naar ervaring van anderen. Prijs-
kwaliteit.","middleexperiencetext":"Nee, ik schrijf nooit reviews.","
simpleratingtext":"Ik zou feedback geven als het in een paar klikken gedaan
 kan zijn. Moet heel simpel zijn, met bv blij/boos emoticons.","
fairratingtext":"Nee, is niet nodig dus ","gamificationrewardtext":"Nee,
niet per se. Zegt niet alles. \nDegene die dat niet hebben, kunnen net
begonnen zijn. Superhosts verhuren waarschijnlijk al langer.","
gamificationchallengetext":"Nee, zeer kleine kans dat je wint.","
gamificationmeaningfultext":"Ja, het is iets concreets wat gebeurd .","
gamificationinteractivitytext":"Ik begrijp deze niet .."}
14 {"gender":"2","age":"56","platformuser":"2","platformusertext":"booking.com\
nuber\nmarktplaats\ntripadvisor\nthe fork\n\n\n","headphonesreviewcontent":
{"gender":"F","name":"Eva van Kampen","picture":true,"review":" I bought my
 pair 1.5 years ago. It feels very comfortable on my head, I can wear it
all day long. I worked in audiotech for a few years, and the guys that are
working on this clearly know what they are doing. This is probably my new
favorite pair of headphones! I added a picture to my review to show what
mine look like."},"headphonesreviewtext":"working in audiotech feels like
this person knows what he is talking about and the picture works for me","
interestreviewtext":"comments that are made by people from who I think did
some research and know something about it","middleexperiencetext":"Yes I
would. Maybe the sellers should write why my opinion is so much worth. \
nNormaly I don't fill in reviews because it is a lot of work and I don't
get the feeling it is worth that much. I only leave a review when I am
motivated\nto warn people away from the product. So explaining why the
review is important works.","simpleratingtext":"Yes it would. Maybe
rankings from 1 - 5 Exellent - to bad. And when you choose a 1 or 5 you can
 explain why..\nSmilies would work for me as well.","fairratingtext":"No, I

would have the feeling I would nog help anyone with another similar review
. It is allready mentioned. So why again.","gamificationrewardtext":"I use
\nstaff traveler. to check what the loads are for a flight. If I give
answers on other peoples requests I get credits to use if I want to know
how full the flight is that I want to take. It works very well. I don't
want to pay for credits so I activly answer requests form other to save
credits","gamificationchallengetext":"Yes, but I would want a sense of how
reliable this company is....are they really giving a free product away? \
n","gamificationmeaningfultext":"Maybe a donation to a good cause and on
the site visible how much money is donated allready. ","
gamificationinteractivitytext":"I don't think so. It sounds like a lot of
work. "}

15 {"gender":"1","age":"24","platformuser":"2","platformusertext":"Uber airBnB
marktplaats thuisbezorgd ","headphonesreviewcontent":{"gender":"F","name":"
Marian Ketelburgh","picture":true,"review":" One month ago I decided to get
new headphones. The rich deep bass works really well for upbeat tracks,
like I'm standing right in the club. I bought something from this brand
before, but was not particularly happy with it. I still wanted to give them
a chance with these headphones, though. I enjoy listening to my favorite
songs with them! You can see what they look like in the picture that I
added!"},"headphonesreviewtext":"Technische details zijn belangrijk voor
mij. ","interestreviewtext":"Of iets van hoge kwaliteit is of niet. ","
middleexperiencetext":"Wiet","simpleratingtext":"Quick 'n dirty","
fairratingtext":"Ja","gamificationrewardtext":"Alleen als een badge
voordelen geeft (kortingen) ","gamificationchallengetext":"Ja, gratis shit
is goed. ","gamificationmeaningfultext":"Pornhub","
gamificationinteractivitytext":"Ja"}

16 {"gender":"2","age":"23","platformuser":"2","platformusertext":"Airbnb","
headphonesreviewcontent":{"gender":"M","name":"Geert Snoeibrug","picture":
true,"review":" One month ago I decided to get new headphones. It feels
very comfortable on my head, I can wear it all day long. I worked in
audiotech for a few years, and the guys that are working on this clearly
know what they are doing. I would recommend this product. You can see what
they look like in the picture that I added!"},"headphonesreviewtext":"
Expert with music","interestreviewtext":"Good use of language","
middleexperiencetext":"If I had the time","simpleratingtext":"Quick way to
give feedback","fairratingtext":"I never write reviews..","
gamificationrewardtext":"Not really","gamificationchallengetext":"Yes bc it
would save money","gamificationmeaningfultext":"Yes this would be a great
trigger ","gamificationinteractivitytext":"Not really"}

17 {"gender":"1","age":"55","platformuser":"2","platformusertext":"Wolwebsite, MP,
YoungLiving oils, Doterra oils, ","headphonesreviewcontent":{"gender":"M",
"name":"Geert Snoeibrug","picture":false,"review":" Last year I purchased
them, and I have used it over the year. The headphones are very durable,

the material feels strong. When I dropped them, there wasn't a single
scratch! This is the first time that I buy headphones from this brand, but
I am sure it won't be the last. I enjoy listening to my favorite songs with
 them!"},"headphonesreviewtext":"Clear, to the point, and also happy with
the material that doesn't fall apart after falling down","
interestreviewtext":"Users: good quality, good material, easy to use, no
tierelantijnen","middleexperiencetext":"I do","simpleratingtext":"A simple
rating is easy, but 2 or 3 open questions are also easy to answer","
fairratingtext":"Maybe, if i have time i do otherwise i forget to do it.\
nMostly i got mail to give a review, when i get the mail and i have time i
do it directly, otherwise i forget it, and don't review anymore","
gamificationrewardtext":"No, i don 't do badges at all","
gamificationchallengetext":"Maybe, if i have time","
gamificationmeaningfultext":"Yes, i go for the tree!\nAnd hopefully i get
the mail when i have time to give the review","
gamificationinteractivitytext":"No, i don 't like that.\nOften it asks for
more time to fill in the options"}

18 {"gender":"1","age":"66","platformuser":"1","platformusertext":"","
headphonesreviewcontent":{"gender":"F","name":"Marian Ketelburgh","picture"
:false,"review":" I bought my pair 1.5 years ago. I love how balanced the
sound of the headphones is, a very natural sound that fits all music. I
worked in audiotech for a few years, and the guys that are working on this
clearly know what they are doing. They do what they must do, which is good.
"},"headphonesreviewtext":"long period of using, user with knowledge of
this type of stuff","interestreviewtext":"review after long period of use",
"middleexperiencetext":"no, too much effort, unless there was some reward",
"simpleratingtext":"a bit, thumbs will do","fairratingtext":"I think I will
 not place a si,ilar review","gamificationrewardtext":"no","
gamificationchallengetext":"yes","gamificationmeaningfultext":"yes","
gamificationinteractivitytext":"no"}

19 {"gender":"2","age":"30","platformuser":"3","platformusertext":"All kinds of
platforms, like Airbnb, Marktplaats, Thuis Bezorgd and a couple platforms
where I follow courses.","headphonesreviewcontent":{"gender":"M","name":"
Geert Snoeibrug","picture":true,"review":" Last year I purchased them, and
I have used it over the year. It feels very comfortable on my head, I can
wear it all day long. I have bought other devices from the same brand, they
 never dissapoint. I would recommend this product. You can see what they
look like in the picture that I added!"},"headphonesreviewtext":"This
review would influence me, I would buy them. T h e r e s a picture, a real
name and this person has used the product for some time.","
interestreviewtext":"If the review is not too old, if the person has
experienced the product for some time and if i t s a real person.","
middleexperiencetext":"Maybe, it depends how easy I can put a review on a
site or platform.","simpleratingtext":"Yes, i t s easy and gives more

information than a thumbs up/down system.","fairratingtext":"I w o u l d n t
have the urge to write something, if t h e r e s a rating system I might
leave my review.","gamificationrewardtext":"I might like extra platform
benefits, but w o u l d n t like a visible badge.","gamificationchallengetext"
:"Yes! Would love this! (I m not sure people will leave honest reviews
though)","gamificationmeaningfultext":"Would really like this, this would
motivate me and people will still leave honest reviews I think.","
gamificationinteractivitytext":"This w o u l d n t motivate me, it only costs
time."}

20 {"gender":"1","age":"25","platformuser":"2","platformusertext":"Thuisbezorgd
(1/2 keer in de maand), airBnb(2-3 keer per jaar), ","
headphonesreviewcontent":{"gender":"F","name":"Marian Ketelburgh","picture"
:false,"review":" I got my fresh pair of Sony headphones last week. I love
how balanced the sound of the headphones is, a very natural sound that fits
 all music. I worked in audiotech for a few years, and the guys that are
working on this clearly know what they are doing. They do what they must do
, which is good."},"headphonesreviewtext":"Positive review by someone who
is an 'expert', or at least experienced and has knowledge about audiotech.
","interestreviewtext":"Depends on the product, but long-term usage/quality
 review of the product, pros/cons, consistency in the reviews (all positive
 or mixed reviews?). ","middleexperiencetext":"Generally yes, but it
depends on the effort to give feedback and the type of product I'm asked to
 review.","simpleratingtext":"A simple rating system would trigger me
faster to leave feedback, then a complex one.\n\nI'm a fan of the rating
system of Medium, the 'claps'. The longer you press the button, the more '
applause' you give. One-sided though, only positive.","fairratingtext":"
Give a rating, maybe. Write a review, probably not. If there are reviews
that already summarize my experience with the product, why write a review.
I'd rather back /rate the review that's close to my experience.","
gamificationrewardtext":"In case of Airbnb, I would trust a 'superhost' as
he/she has got positively reviewed.","gamificationchallengetext":"If the
product has value, I would consider it. I think it would trigger some
dishonest, or overly extensive, reviews.","gamificationmeaningfultext":"Yes
, I would write a well-written review. But I would demand more info (how,
where, when, progress) on the 'planting tree' promise.","
gamificationinteractivitytext":"Yes, backing up other reviews is 'low
effort'. Would do that."}

21 {"gender":"1","age":"25","platformuser":"3","platformusertext":"FB: daily\
ninsta: daily\ntinder: daily\nbumble: daily\nthuisbezorgd: zelden\
ngreenwheels: ongeveer 2x per maand\nudemy: soms\nreddit: daily\n\n","
headphonesreviewcontent":{"gender":"F","name":"Eva van Kampen","picture":
true,"review":" One month ago I decided to get new headphones. I love how
balanced the sound of the headphones is, a very natural sound that fits all
 music. This is the first time that I buy headphones from this brand, but I

am sure it won't be the last. I would recommend this product. I added a picture to my review to show what mine look like."},"headphonesreviewtext": "also 1 cause its a bad review. but the overall reviews were good. also this one had a picture.","interestreviewtext":"credibility, if the writer uses proper language. if the writer of the review doesnt gain something from giving it a higher rating. \namount of reviews. two 4 star reviews is more believable than one 5 star review.","middleexperiencetext":"if i had incentive. but i never bought something and then went back to the place i bought it to leave a review\n","simpleratingtext":"probably. * out of five works for me. ","fairratingtext":"maybe a short one. might be nice to have a quick \"seconded\" option or something. where you can back someones response ","gamificationrewardtext":"yes. if people earned badges by being active and leaving proper reviews. thats makes them more reliable. \nairbnb superhost's are immediately more credibly.","gamificationchallengetext":" nah, to much incentive to write to much. or exaggerate. cant have people making t much of an effort.","gamificationmeaningfultext":"nice! i would like that. would for sure try to write good reviews then."," gamificationinteractivitytext":"nah things need to be clear at a glance. "}

22 {"gender":"2","age":"24","platformuser":"2","platformusertext":"Airbnb, 3 times a year I use it often (3 times a day for 3 days approx.)"," headphonesreviewcontent":{"gender":"M","name":"Ronald Jonkheer","picture": false,"review":" 3 days ago I got my hands on these headphones. It feels very comfortable on my head, I can wear it all day long. This is the first time that I buy headphones from this brand, but I am sure it won't be the last. I enjoy listening to my favorite songs with them!"}," headphonesreviewtext":"Good value for my money and the reviewer praises the quality of the product instead of looks (like fashionable and scratches)", "interestreviewtext":"Value for money, durability, explanation on why they feel this way. But also feeling, what feels to be the truth"," middleexperiencetext":"I'm always planning on doing it but I never take the time. ","simpleratingtext":"I like stars the most because it says more than just a (dis)like) it's easier to claim that you love the product except for a few minor details that annoy the heck out of you. And thus you give it 3 stars. ","fairratingtext":"Probably not","gamificationrewardtext ":"Probably not, It seems like it is too much work to get that so I would not put in the effort because it feels like a bigger commitment to reach a badge. ","gamificationchallengetext":"Maybe. It would stimulate me to think more about it but the changes are that I open the site, leave it open for a few days and still don't write that review. ","gamificationmeaningfultext ":"Same as the previous answer: (Maybe. It would stimulate me to think more about it but the changes are that I open the site, leave it open for a few days and still don't write that review.)","gamificationinteractivitytext": "No probably not. Because I would not be online enough on this specific site to see the customization options daily. "}

23 {"gender":"1","age":"25","platformuser":"3","platformusertext":"Uber: rarily on holiday\nGrab: on holiday (in Asia)\nMarktplaats: occasionally\nOnline shopping platforms bol.com/coolblue/aliexpress etc regularly","headphonesreviewcontent":{"gender":"M","name":"Bart Ochtendschalker","picture":false,"review":" I bought my pair 1.5 years ago. The headphones are very durable, the material feels strong. When I dropped them, there wasn't a single scratch! I worked in audiotech for a few years, and the guys that are working on this clearly know what they are doing. I would recommend this product."},"headphonesreviewtext":"The guy has experience with audio and has been using the headphones for quite some time already.", "interestreviewtext":"An honest opinion of someone who used the device for some time. Someone who seems reasonable and took a little bit of time writing the review.","middleexperiencetext":"Yes maybe, often I don't care to leave a review if I am just fine with the product. Only if it is bad I might review, very rarely if it is very good. I think the middle reviews disappear because platforms tend to show only the good reviews (at least at plain sight).","simpleratingtext":"If it is very easy and quick I will be more inclined to 'leave a review'.","fairratingtext":"I don't think so, since my opinion is already available to potential buyers. Maybe I would 'plus' or 'thumbs-up' the review if possible.","gamificationrewardtext":" Maybe it would incentivize if i am already present on the platform, but it would not convice me to get started in the first place.","gamificationchallengetext":"Potentially, if the pool of potential reviewers is not too big, otherwise the chance of me winning is too slim.","gamificationmeaningfultext":"No, I don't see the tree being planted so I am a bit sceptical if that will happen at all.","gamificationinteractivitytext":"Not really, I don't really care about my digital (anonymous) profile."}
24 {"gender":"2","age":"23","platformuser":"2","platformusertext":"Airbnb, Marktplaats, couchsurfing, UW ","headphonesreviewcontent":{"gender":"M"," name":"Bart Ochtendschalker","picture":false,"review":" Last year I purchased them, and I have used it over the year. I love how balanced the sound of the headphones is, a very natural sound that fits all music. I have bought other devices from the same brand, they never dissapoint. I would buy this again if it ever breaks."},"headphonesreviewtext":"This Bart used the headphone for a year which is more trustworthy then a review of a 1-month-user. Besides he's the (almost) only one focussing on how the sound of the headphones actually is. ","interestreviewtext":"I first read the most negativ reviews and check if the amount of them is enough to trust it. (If its 2 negativ reviews against 20 positiv reviews I would give the product a chance) Another thing I would look at is which reviewer gives a really detailed description of the product. ","middleexperiencetext":"Only if it's from a smaller company (to support them), bigger companies are also selling lots of stuff withouth me writing a review. ","simpleratingtext":"

Maybe but i find less detailed reviews quite useless so when i see only a star-rating it's not telling me anything about why somebody gave 1 star. \nIdeal: probably start with star-rating and then when i click on a star (for example 2 stars) i would see the reviews of people who chose and explain their 2 star-rating. In that way it's clear and orderly. ","fairratingtext":"No, i would 'like' the already existing review when possible. ","gamificationrewardtext":"No it triggers a highly competitive surrounding which i don't like. It almost forces u to also try to become a, in this example, superhost, which means you're behind when you dont put so much time and effort in it. ","gamificationchallengetext":"Again a bit competitive, besides, companies already sent professional reviewers free stuff so they can write a real long, detailed review about their product. ","gamificationmeaningfultext":"YES i would defenitely want to write. BUT i hope the company will also plant a tree without me writing a review, otherwise they would put the responsibility for taking care of our planet on their buyers/users. ","gamificationinteractivitytext":"Maybe, but by using the word 'rate product features' i realise i would be triggered to rate a product when the website already gives me multiple product related themes to rate. (Such as 'sound', 'material', 'volume', 'sustainable', 'easthetics' for a headphone product) "}

25 {"gender":"2","age":"24","platformuser":"2","platformusertext":"AirBnB, marktplaats and Uber","headphonesreviewcontent":{"gender":"M","name":"Geert Snoeibrug","picture":false,"review":" One month ago I decided to get new headphones. The battery on these headphones is amazing, they last one week of daily use on a single charge. I worked in audiotech for a few years, and the guys that are working on this clearly know what they are doing. I enjoy listening to my favorite songs with them!"},"headphonesreviewtext":" Because it mentions a subject that I think is important when looking for a wireless headphone (the battery), and Geert says he worked in audiotech. So that sounds like he knows what he is doing.","interestreviewtext":"I want to know if the product is different than how the website describes it, and if it lasts long enough to be worth the price.","middleexperiencetext":"Not sure, because that would mean that you should review ALL of your purchases and giving a middle review seems awkward. I would not know what to write:\n\"This headphone was rather normal. You should buy it if you're not looking for something extraordinary.\"\n\"This headphone is the most average product you can get, 3/5.\"","simpleratingtext":"Yes, when it doesn't require writing a text rating is easy. \nI like the 5 star rating systems. ","fairratingtext":"No, I would only write a review if I feel like I can add some new information on the website.","gamificationrewardtext":"If it would give benefits (not only a visible badge) it would give me motivation to leave a review, because it seems little effort for nice benefits.","gamificationchallengetext":"This would not motivate me, because the chance of writing the best review seems not that high. Especially

because there are people who write reviews as a profession. Also, if the
current product works fine, why would I want the next edition too? Is it
already broken by the time a new edition comes out?\nAnd I wonder if the
would only select people who write a positive review about the product. Who
selects \"the best review\"?","gamificationmeaningfultext":"I think this
would be a nice motivation, but it would only work ones of twice. It's the
same as that search engine that plants trees. You try it a few times and
then get back to your old habits of using Google.","
gamificationinteractivitytext":"I do not see the direct benefit of giving a
review in this situation, so I don't think it will motivate me."}

26 {"gender":"1","age":"32","platformuser":"2","platformusertext":"Airbnb,
Marktplaats, Werkspot, LinkedIn (for job search), ","
headphonesreviewcontent":{"gender":"M","name":"Geert Snoeibrug","picture":
false,"review":" 3 days ago I got my hands on these headphones. The battery
on these headphones is amazing, they last one week of daily use on a
single charge. I have bought other devices from the same brand, they never
dissapoint. They do what they must do, which is good."},"
headphonesreviewtext":"Review 6 (Geert Snoeiburg), together with review 2 (
Bart Ochtendschalker) influence me to buy the headphones, because they both
give some technical details which I care about, and they seem like they
know what they are talking about.\nThen there is one negative review, but
for me that is royally canceled out by the other 3 remaining reviews which
are positive. ","interestreviewtext":"Normally I don't let 1 review
influence me, I look at them all and try to include all the reviews which
seem genuine before I make a decision. (I try to ignore reviews which seam
fake or not useful). Also, I almost always look for several competing
products and compare the reviews of those different products to each other.
","middleexperiencetext":"Maybe, but I almost never write any reviews, also
no 'upper' or 'lower'-experience reviews.","simpleratingtext":"I do rate
movies/series at iMDB, I rate (star) all the movies there that I have seen,
but that is also because in that way I can always see if I have already
seen that movie or not. (I don't write reviews with text there either). I
use the thumbs up / down button in YouTube in a similar matter: I like
movies that I want to see again. I almost never use the thumbs down button,
only sometimes when I really don't like what I see.","fairratingtext":"
Probably not. I almost never write reviews, and when I think the current
rating is accurate, I will probably not write a review.","
gamificationrewardtext":"Yeah maybe, depending on what the benefits would
be and how much I would have to do for that. But a bit of community
building is nice of course! And hats. I like hats. Can I get a superdude
hat?","gamificationchallengetext":"Yes, definitely. This is a great idea.",
"gamificationmeaningfultext":"Wow! Maybe I will write some 'fake' but well
written reviews now and then, or maybe I will even try to write a script to
generate tons of reviews! Nice.","gamificationinteractivitytext":"Sounds

good, I would love to rate certain aspects of a product in stead of one overall score. It would make it easier, but only when I can decide for myself how many aspects I want to give a rating to."}

27 {"gender":"2","age":"25","platformuser":"3","platformusertext":"Airbnb, once or twice a year. Indigogo, once a year. Kiva, once a month.","headphonesreviewcontent":{"gender":"M","name":"Bart Ochtendschalker","picture":false,"review":" Last year I purchased them, and I have used it over the year. The headphones are very durable, the material feels strong. When I dropped them, there wasn't a single scratch! I worked in audiotech for a few years, and the guys that are working on this clearly know what they are doing. They do what they must do, which is good."},"headphonesreviewtext":"The guy says he has experience with audiotech and experience of a year with the headphones.","interestreviewtext":"If there aren't any \"traps\" in buying the item. Am I forgetting to look at an important spec or aspect of the item?","middleexperiencetext":"I think I still wouldn't. I don't give reviews in general, often, in my opinion, there are already enough reviews for many things. I only gave a positive review once, when I felt connected to the maker of the product. In your explanation you mention that feedback is important for people. However, when giving reviews, it is like giving reviews to a company, a big black box, not to people. That is why emotionally it doesn't feel important. ","simpleratingtext":"I think on AirBnB I have been persuaded to rate three factors on a scale from 0 to 10. This would be on the points of cleanliness , good contact with the owner ect. I think this would be my ideal rating system. Just three clicks, when distributed over a lot of customers more than 3 questions could be answered. Everyone would just get three random questions assigned. This way the burden is low, making it more likely that many people respond. A 10 point scale is in my opinion more likely to not receive a U-shape. I am making the assumption here that people associate stars with awards, each star must be earned. Whereas on a 10 point scale is more associated with high school tests, which give nuance in great, good enough or not good enough.","fairratingtext":"No. But that is also related to my answer on question one.","gamificationrewardtext":"It depends on the system. Getting a reward for rating a certain number of products does not motivate me. On Duolingo though the rewards do help, as they help me get more motivated for a goal I am already motivated for.","gamificationchallengetext":"I think a good review is subjective. A free product is a good reward, however if the product has bad quality does this make people send in illegitimate good reviews?","gamificationmeaningfultext":"I would love this! But that is also because I love trees. I don't know how big the tree-loving community is, and if this would skew reviews in the direction of tree-positive reviews. (Tree hating people might not be motivated by this)","gamificationinteractivitytext":"Once again, I don't feel motivation here, leaving a review still feels like leaving a review

```
for a company. Making the review system more complicated would not change
this."}
```