Mikko Pirilä

DON'T FEED THE TROLL:

RESPONSE STRATEGIES FOR FALSE REVIEWS AMONG SMALL FINNISH SERVICE SECTOR COMPANIES

ABSTRACT

Pirilä, Mikko
Don't feed the troll: response strategies for false reviews among small Finnish service sector companies
Jyväskylä: University of Jyväskylä, 2022, 91 pp.
Information Systems, Master's Thesis
Supervisor: Soliman, Wael

As online reviews are only second to friends and family and read regularly by more and more consumers, their impact on purchase decision is considerable. A byproduct of this, false reviews, are reviews that deliberately aim to either boost or defame their target, such as a restaurant or a hotel via praising, or slanderous messages. Considering that false reviews have been observed to be difficult for humans to detect, it is obvious that a company without a strategy for dealing with fake reviews could be in big trouble, if they are targeted by such a campaign.

The empirical part of this thesis looks at false review from the perspective of companies and how they are suited to handle them. Also, the interviewees deterrents on purchasing false reviews are explored. The security action cycle, which consists of four phases: deterrence, prevention, detection, and remedies is used to guide the empirical part, which is conducted as semi-structured interviews. A total of eight participants from small Finnish service industry companies are interviewed on their stance on false reviews in order to answer the research questions:

1. What actions are companies employing to prevent damages from fake internet reviews?

2. What factors deter a company from colluding with fake reviews

The data gathered from the interviews is then thematically analyzed and presented in relation to the security action cycle phases. The major findings point that the preparations are generally lacking due to the absence of any serious cases of false reviews in Finland. As the false reviewer was mainly imagined as a disgruntled individual, most respondents would just disregard it with the occasional response to the review as a remedy. In a case of more serious situation, most would contact the platform for example, but the hopes of it amounting to something were low.

Also, the general attitude towards buying reviews was negative and more so towards the negative false reviews. The deterrents mentioned by the interviewees were related mainly to ethics and getting caught and its consequences. However, the more entrepreneurial individuals saw buying positive reviews akin to buying advertisements instead of condemning them outright.

Keywords: False reviews, Security action cycle, Online reviews, Buying reviews, Deterrence

TIIVISTELMÄ

Pirilä, Mikko
Don't feed the troll: response strategies for false reviews among small Finnish service sector companies
Jyväskylä: Jyväskylän yliopisto, 2022, 91 s.
Tietojärjestelmätiede, pro gradu -tutkielma
Ohjaaja: Soliman, Wael

Internet -arvostelut jäävät toiseksi vain perheen ja ystävien suosituksille ja että niitä luetaan yhä enemmän, on niiden vaikutus ostopäätökseen huomattava. Internet -arvosteluiden sivutuote, valearvostelut ovat arvosteluita, joiden tarkoitus on ylistää tai halventaa arvostelun kohdetta, kuten esimerkiksi ravintolaa tai hotellia. Koska valearvosteluiden on todettu olevan vaikeasti havaittavia ihmisille, on selvää, että yritys, jolla ei ole suunnitelmaa valearvosteluiden varalle, voi joutua suuriin vaikeuksiin joutuessaan valearvostelukampanjan uhriksi.

Tutkielmaa ohjaamaan käytettiin teoriaa turvatoimien syklistä *(engl. Security action cycle)*, joka koostuu neljästä vaiheesta: pelotteet, esteet, havaitseminen ja korjauskeinot. Tutkimus toteutettiin puolistrukturoituna haastatteluina.

Tutkielman empiirisessä osuudessa valearvosteluita tarkastellaan yritysten varautumisen näkökulmasta. Myös haastateltavien pelotteita valearvosteluiden ostamista kohtaan tutkittiin tarkemmin. Haastattelemalla kahdeksaa osallistujaa pienistä suomalaisista palvelualan yrityksistä, pyrittiin saamaan vastaukset tutkielman tutkimuskysymyksiin:

1. **Mitä toimia yrityksillä on käytössä torjuakseen valearvosteluista syntyviä haittoja?**
2. **Mitkä tekijät toimivat pelotteena valearvosteluiden ostamiselle?**

Haastatteluista kerätty data analysoitiin temaattisesti ja esitettiin suhteutettuna turvatoimien syklin teorian eri vaiheisiin. Tutkielman merkittävimmät löydökset liittyvät näiden torjumiskeinojen puutteeseen, joka puolestaan haastateltavien mukaan johtui vakavampien valearvostelutapausten puutteesta suomessa. Koska valearvostelija nähtiin yleisesti ärsyyntyneenä asiakkaana, suurin osa vastaajista päätti olla välittämättä, kuitenkin satunnaista arvosteluun vastaamista lukuun ottamatta. Vakavammissa tapauksissa vastaajat mm. ottaisivat yhteyttä arvostelualustoihin, mutta toivot sen toimimisesta olivat vähäiset.

Yleinen mielipide valearvosteluiden ostamista kohtaan oli negatiivinen. Positiivisten arvosteluiden ostamista ei kuitenkaan nähty yhtä pahana kuin negatiivisten arvosteluiden ostamista. Näihin liittyvät pelotteet liittyivät pääosin etiikkaan ja kiinnijäämiseen ja sen seuraamuksiin. Kävi kuitenkin ilmi, että yrittäjähenkisemmät yksilöt näkivät positiivisten arvosteluiden ostamisen vähemmän pahana asiana, sillä he rinnastivat sen mainoksiin ja niiden ostamiseen.

Asiasanat: Valearvostelut, Turvatoimien sykli, Internet -arvostelut, Arvosteluiden ostaminen, Pelote

FIGURES

TABLES

SISÄLLYS

1   INTRODUCTION

This thesis looks at the phenomenon of false reviews on internet review platforms and other social media from the viewpoint of how companies see them and how they are prepared to combat the potential harm related to them. Such platforms are for example google reviews, TripAdvisor, and Yelp, whereas the social media in this thesis is mainly limited to Facebook and the reviews within.

It is important to bring light into what the companies are doing regarding false reviews as the effect of reviews on purchase intentions has been proven time and time again (Mathwick & Mosteller, 2016); (Sutanto & Aprianingsih, 2016). False reviews on the other hand have been proven to damage the involved business's reputation (Ahmad & Sun, 2018) and the customers knowledge of them has been observed to lower purchase intentions (Munzel, 2016). The false reviews are also deceptively common occurrence, as the review platform Yelp.com suspected at least 20% of its comment of being fake (Munzel, 2016). Sometimes referred to as the eWOM, or electronic word of mouth, the online reviews are only second to the recommendations of family and friends (Mathews Hunt, 2015).

The empirical research done in this thesis is conducted as semi-structured interviews via Zoom and in some cases by phone. The sample for the study consisted of eight case companies, with the interviewees being at low to top management levels in different small service sector companies. Small companies were specifically chosen, as in general, getting reviews is good for a company and for smaller, more unknown companies, the effect from online reviews is stronger while large chain affiliated companies are less affected by them due to the extensive marketing campaigns, they have conducted to boost their visibility (Luca, 2011). The interviews are then thematically analyzed and presented in order to answer the research questions:

1. What actions are companies employing to prevent damages from fake internet reviews?

2. What factors deter a company from colluding with fake reviews

In general, the companies are well prepared for what the current situation regarding false reviews is. As the current situation consists mainly of individual disgruntled customers, the defenses can afford to be lax. The defenses consisted

mainly of ignoring the suspected false reviews due to the minimal impact of an individual review, with the odd response here and there. Most also thought that their pre-built reputation would deter the false reviewer in the first place, as well as make the false review have little impact for example. However, in the case of a more serious campaign many of these establishments could face serious challenges due to the lack of planning regarding false reviews.

As for the buying of false reviews, the main deterrents were found to be ethics and the fear of consequences. Also, buying negative false reviews was condemned more harshly than buying their positive counterparts. Some even equated the positive false reviews to advertising.

The literature review first looks at the phenomenon of online reviews in general. This first chapter of literature review goes over general matters related to online reviews, such as the effects of reviews and review platforms. Only after online reviews have been introduced sufficiently, the false reviews can be discussed. The second chapter of the literature review focuses on these falsified online reviews, for example by discussing their effects, spread and reasons for existing. After the literature review, the security action cycle by Straub and Welke, (1998) is introduced, which is used to guide the thesis. This cycle consists of four stages: deterrence, prevention, detection, and remedies.

After this, the research method is introduced. The research method chapter goes over the research design, analysis methods, the questions used to guide the interviews, and the chosen sample for the study. The findings are then presented in relation to the security action cycle phases, which then provides the answers to the research questions of this thesis.

Due to their prevalence and great importance for businesses, the online reviews have been extensively studied outside of academic literature. For the purposes of this thesis, surveys and other industry reports are used to complement the academic literature with additional statistics for example. An example of surveys used in this thesis is the Local Consumer survey by Brightlocal (Murphy, 2020), which is based on a representative sample of US-based consumers with n=1013.

2   LITERATURE REVIEW

This chapter introduces online reviews and false online reviews with the different factors related to them based on previous academic literature. The first chapter, which introduces the online reviews, brings up the effects of online reviews in general, as well as what makes them reliable for example. Also, the platforms hosting these reviews are introduced at sufficient depth. After laying the foundation by properly introducing the internet reviews in general, the fake ones are introduced. The second chapter of the literature review looks at these falsified internet reviews and what their effects are, how they spread and what is typical of them. Also, the buyers and sellers of the reviews are discussed.

2.1   Online consumer reviews and electronic word of mouth

With the traditional word of mouth no longer being the only way for consumers to spread opinions about products and services due to the popularity of reviewing platforms and social media, there is more information involved in the purchase decision process than ever before. These online reviews, commonly referred to as eWOM (electronic word of mouth) or OCR (online consumer reviews) are only second to the suggestions of family and friends, beating all other traditional forms of advertising (Mathews Hunt, 2015). Therefore, it is not surprising that the effect of online reviews on purchase intention is great.

These reviews are written and read by consumers in order to share experiences related to the product or service at hand. Whether it is on the products sales page, a review platform or on a Facebook page/group related to the product or service, the online reviews and electronic word of mouth influence their readers. All implied numerical ratings in this thesis are on a scale of 1 to 5 with 1 being the worst and 5 being the best rating unless explicitly stated otherwise.

The broader term often used in the context of online reviews, electronic word of mouth (eWOM), which covers the concept of OCR. Hennig-Thurau, Gwinner, Walsh and Gremler, (2004) define eWOM as *"any positive or negative*

statement made by potential, actual, or former customers about a product or company, which is made available to a multitude of people and institutions via the Internet.", while online consumer reviews (OCR) are generally defined in academic literature as the opinions of consumers about a service or a product that they have previously purchased (Zhang, Cheung & Lee, 2014).

2.1.1 The effect of online reviews

Whether a review is positive, neutral, or negative, it affects the reviewed companies' reputation. While positive reviews towards a company increase its outlook in the eyes of the consumer, the negative reviews may do the opposite. However, getting reviews in general seems to be a net positive for a company, as per study by Vermeulen and Seegers, (2009), who studied hotel reviews and found out that consumers are more aware of them if they have more reviews in general, with positive reviews additionally improving the consumers' attitudes towards the hotels. This is also supported by Motoyama and Usher, (2020), who found that on average, as a restaurant gains more reviews, the average rating goes up. Vermeulen and Seegers, (2009) also noted that the online reviews had a stronger effect on lesser-known hotels, which is in line with Luca, (2011), according to whom the impact of online review websites is more substantial on lesser-known establishments in a study about Yelp reviews. In the same study Luca, (2011) found that chain restaurants do not seem to be affected by the Yelp ratings as they have an established reputation built through extensive marketing and advertising.

2.1.2 Purchase intention

The effect of online reviews on customer purchase intention has also been studied broadly. For example, Sutanto and Aprianingsih, (2016) found a positive relationship between purchase intention and the following factors of online reviews: source credibility, review quality, review quantity and review valence. On the other hand, a slightly conflicting discovery was made by Maslowska, Malthouse and Bernritter, (2016) that suggests that an increase in the review volume does not always equate to more sales. However, the studies were conducted on different types of products, which may indicate that different products or demographics benefit more from the volume of reviews for example.

According to a survey by PowerReviews (O'Neil, 2015), 95% of consumers regularly read reviews when researching a product, they intend to purchase. Lackermair, Kailer and Kanmaz, (2013) had similar results, when measuring online review utilization by consumers, ending up with around 85% of people reading reviews often or very often when buying products online. In the same study, 74% of people thought that online consumer reviews played an important or very important part in the purchase intention (Lackermair et al., 2013). Also, the higher the initial interest in the product is, the more reviews and more carefully a prospecting customer will read in order to make their decision (Filieri, 2016).

Having a high review score makes a business more attractive in the eyes of the potential customer and thus leading to higher revenue for the business. This has been proven for example by Luca, (2011), who in his study saw that an increase of one star (out of five) for independent restaurants on Yelp.com equaled to a 5-9% increase in revenue. Also, on the other end of the spectrum, having a low review score means that potential customers may shy away from the business because of it. Additionally, according to a survey by Brightlocal (Murphy, 2020) 52% of consumers would not consider using a business with a rating less than 4 out of five and 81% of consumers would not using a business with a star rating of 3 or less.

The effect of word of mouth on purchase intention is also mediated by the consumers initial attitude and views about the product or service when the brand in question is familiar (East, Hammond & Lomax, 2008). East et al., (2008) found out that negative word of mouth on a brand that a consumer prefers are resisted, while positive word of mouth on brands that a consumer does not prefer are also resisted. They also note that the impact both negative and positive word of mouth is closely tied to the initial purchase intention prior to the word of mouth, making an impression that consumers tend to read the online reviews in order to reinforce their own views to a certain extent.

2.1.3 Positive versus negative reviews

Negative reviews are generally seen as more trustworthy than positive and neutral ones (Ullrich & Brunner, 2015; Baek, Ahn & Choi, 2012). The weight of the negative reviews is increased in unfamiliar brands while familiar brands did not experience any significant differences in terms of weight (Ahluwalia, 2002). Thus, it can be assumed that a positive review of the same "intensity" as a negative one is less influential and cannot undo the harm done by the negative review. However, the positive reviews seem to be more influential when it comes to the consumers purchase intention, at least in the case of familiar brands according to East et al., (2008).

While every company wants their products and services to be rated highly, the absence of negative reviews can undermine the sales effort and make the company seem suspicious. Negative reviews have been found to legitimize reviews, as the credibility of a review suffers if there are no negative reviews at all while the overall rating is very high (Doh & Hwang, 2009). Attempting to remove these negative reviews with the intention to preserve a "perfect image" can also cause companies more harm than good, as a few negative reviews are not critical to the rating if the other reviews are positive (Doh & Hwang, 2009). This is confirmed by Maslowska et al., (2016) who found out that the feeling of "too good to be true" seems to be present for the readers of reviews, as products in the rating range of 3.8-4.2 (out of 5) were more likely to be bought than those with very high ratings. They suggest that customers often get the feeling of foul play when dealing with very highly rated products.

2.1.4   The early reviews

Due to the nature of the reviews, they are posted as more and more consumers experience the service or product. This in turn cause the reviews that exist for the longest, to affect the consumers' opinions the longest. The early reviews are extremely important for companies, as they tend to dominate the helpfulness metric on the review websites (Lu, Wu & Tseng, 2018), making them "weigh" more in the eyes of the consumer. Considering this "pioneering effect", the initial reviews a company gets, need to be managed properly, lest the reviews take a wrong direction from the get-go.

An important factor to consider in relation to the initial reviews is that the innovators and early users' opinion can also vary greatly from the opinions of the majority, which may warp the "real" rating (Li & Hitt 2008). This in turn can be a make or break for the company as the reviews that the majority will read are the early ones posted by the early users and innovators which set the tone for the product or service. The difference between the early users and the majority can manifest itself as a bimodal review graph, meaning that the reviews graph has two peaks, typically one that is very high and one that is very low, due to the conflicting opinions of the different user groups. This phenomenon is discussed in more depth in a later chapter.

Even though the early reviews are usually pinned as the most helpful ones, in general consumers tend to disregard older reviews and only focus on the more recent ones. A result from a survey by Brightlocal (Murphy, 2020) is that 73% of consumers tend to disregard reviews that are older than a month. However, it is to be noted that while services such as restaurants might have a decrease or an increase in quality over time, making recent reviews the most trustworthy. Products such as electronics on the other hand do not, assuming they are sold as new and are the same version etc., thus it could be assumed that this effect is mainly present in services where the quality can change.

It should be of utmost importance for a new company to get reviews, as a company with only a few reviews is likely to have a lower score (Motoyama & Usher, 2020). Motoyama and Usher, (2020) suggest tackling this by only showing a company's rating after a certain number of reviews have been received, as to thwart biased opinions based on the low number of reviews.

As most people utilize online reviews in their purchase process, a company cannot overlook the effect that online reviews have on their business. To take the full advantage and prevent backlash from the reviews, they must be managed. The next chapter looks at what action's companies can take to manage their reviews online and what effects they may have.

2.1.5   Review management

A well-managed review section can be an ample opportunity for a company to co-create value with its customers. Exchanging information between the company and the consumers allows for the companies to get valuable feedback on what they have succeeded in and what needs improvement. Likewise, from the

consumer's point of view, the reviews are one of the most important ways of getting unbiased information about a company and its services. The term value co-creation refers to the customer being an active participant in the value creation process (Vargo & Lusch, 2004), which in the online review process is the provider of value to the company and to the other consumers.

A company facing bad reviews, or even false reviews, can lessen the damage done by managing their reviews and actively interacting with them. By challenging the negative reviews with tailored messages, a company can impact the subsequent opinions of the said negative reviews positively (Wang & Chaudhry, 2018), however, a positive review counteracts the negative ones more effectively than a brand response (Ullrich & Brunner, 2015). Shin, Perdue and Pandelaere, (2019) also note that the personalized messages have a higher impact on negative reviews compared to positive ones. From the consumers' perspective, this may imply that the company has acknowledged its shortcomings and will try to fix them. However, Wang and Chaudhry, (2018) also found out in their study that highlighting a positive aspect of a review had a negative impact as it was seen as taking an advantage of the positive message by the interviewees. As consumers tend to compare the positive and negative reviews quite often, (83% of surveyed participants) intention (Lackermair et al., 2013), the responses to the negative reviews are likely to be noticed and compared to the other reviews.

These responses have different effectiveness depending on the company in question, according to Ullrich and Brunner, (2015), weak (relatively unknown) brands' responses on negative reviews have a lesser effect compared to their well-known counterparts who have their brand as assurance of quality.

As getting reviews is net positive for a company, the logical goal is to get as many reviews as possible. Personalized messages are a good way to engage with the consumers as they feel more empowered to co-create information value after they read or receive personalized messages on their reviews (Shin et al., 2019). Confronting the false reviews as a countermeasure based on Lappas, Sabnis and Valkanas, (2016) work will be looked at in more depth later in this thesis.

2.1.6   Review platforms

There are three different types of websites that users can write reviews on: *"independent consumer review websites (or online community), third-party e-commerce websites, and corporate websites"* (Filieri, 2016).

The reviews on specific products are typically located on the seller's website or the third-party e-commerce websites, while bigger review platforms and independent consumer review websites, such as yelp.com and tripadvisor.com tend to review the businesses themselves. The review platforms are especially useful in cases where the business offers services and doesn't necessarily have a review option in place on their website for example. Also, these larger review platforms, such as TripAdvisor have defenses in place to combat false reviews through artificial intelligence and flagging of comments for further inspection by a moderator for example ("What does Tripadvisor do about unfair reviews?" (n.d.)). Also, the wide array of punishments that the platforms can impose on the

false reviewers and those found buying them for example deter the possible offenders from trying to false review in the first place (Straub & Welke, 1998). These platforms also prevent the reviewed companies from removing negative reviews, making them somewhat more reliable than review sections controlled by the product/service provider. For example, Tripadvisor only removes reviews if they are found to violate their guidelines ("What does Tripadvisor do about unfair reviews?" (n.d.)).

The reviews on the different review platforms and product pages are not the only places to find information on a company for the consumer. As the term electronic word of mouth suggests, basically any mention of the company on the internet can be considered to affect the reputation of the company. Online forums, instant messengers, blogs, chat rooms etc. can also be influential in affecting the reputations of companies. For example, many companies have twitter accounts, Facebook groups among others for social media visibility purposes and marketing, where people can comment and share opinions on the services and products. A recent example of attempting to affect reputation on an online forum is the 2013 Samsung false review scandal, where they were found to have hired writers to fabricate negative messages about their competitor on a Taiwanese online tech forum, as reported by the Verge among others (Souppouris, 2021). Also, the different social media pages and official websites that a company has, are typically overseen by the company, giving them near full power to control what types of comments and other contents are allowed, which in turn makes them less trustworthy (Filieri, 2016). Messages posted by the companies' different managers are generally seen as less trustworthy than those of an actual customer (Sparks, Perkins & Buckley, 2013; Mathews Hunt, 2015).

2.1.7 What makes a review seem trustworthy

As mentioned by Mathews Hunt, (2015) online reviews beat the traditional advertising strategies in terms of trust and reliance and are only second to friends and family. With most consumers (Lackermair et al., 2013) utilizing online reviews when researching products and businesses, the general trust towards online reviews can be considered to be quite high. However, when assessing individual reviews, several factors related to the review, reviewer, and the reader themselves, affect whether the review is seen as legitimate. This in turn has an effect on the purchase intention as the higher the perceived trustworthiness of a review is, the higher effect it has on the readers decision making while the untrustworthy reviews are discounted from the decision-making process (Filieri, 2016).

The following sections explain a theoretical framework by Filieri, (2016) which is then compared to other academic literature to complement the results. The framework suggests that there are four review-based factors: source trustworthiness, message trustworthiness, review valence, and pattern in reviews. Also, the medium in which the reviews are posted on affect the perception of trust. These categories then contain several different factors contributing to them. A reader's assessment of source trustworthiness is also affected by their

involvement with the product/service in question and their prior experience with online consumer reviews. (Filieri, 2016)

2.1.8 Source trustworthiness

- Source trustworthiness which is based on examining the reviewer's profile.
    - Consumers experience with ORCs in general assists in determining review trustworthiness.
    - Consumers who are highly involved with a product or service read OCRs more carefully to assess the trustworthiness than those with low involvement
    - The more contributions and the more information that a reviewer discloses about themselves, the more trustworthy he or she is in the eyes of the reader.
    - Also, how many reviews the user has posted, how extreme they are, the valence of the ratings, and the content of the messages affect trustworthiness.
    - For example, one-time reviewers with an extreme rating are in general considered untrustworthy by the readers.
    - A given reviewer is deemed more trustworthy by the reader if they disclose information about themselves, such as a picture or a bio text, than a reviewer with no information about themselves (Munzel, 2016). Also, a verified buyer badge has been proven to increase review credibility and purchase intention consequentially (Kim, Maslowska & Malthouse, 2017)

However, according to Martens and Maalej, (2019) false reviewers on the Apple Appstore tend to have up to 12 more reviews than the legitimate reviewers, which can play into the false reviewer's favor as they are deemed more reliable due to the number of reviews.

2.1.9 Message trustworthiness

- Message trustworthiness depends on its length, type of information, detail, style, and included pictures.
    - Longer reviews are considered more trustworthy
    - Length in itself is not enough though, as a review must have factual, detailed and relevant in order to be trustworthy.
    - Emotional and subjective messages are seen as untrustworthy
    - Included pictures of the purchased product increase the trustworthiness of a review greatly.

The perceived information quality positively affects trustworthiness in online reviews and does not depend on the review rating, while poor information quality decreases trust towards the review (Zhu, Li, Wang, He & Tian, 2020).

2.1.10 Review valence

- The review valence i.e., the rating, affects the perception of depending on if it is negative or positive.
  - Negative reviews are considered more trustworthy, which is the general consensus among relevant literature as well (Ullrich & Brunner, 2015; Baek et al., 2012) for example.
  - Extreme reviews (1 or 5) are seen as less trustworthy as they are more likely to be the product of review manipulation than moderate ratings.
  - Reviews, which discuss both negative and positive factors are seen as more trustworthy.
  - The readers may get a feeling of "too good to be true" if the review score is very high and lacks negative ratings Maslowska et al., (2016).

2.1.11 Patterns in reviews and the medium

- Highly involved customers can start to notice a pattern in the reviews.
  - Many great reviews posted by one-time reviewers, many of which are typically posted during a short timeframe
  - A consensus among the reviews increases trustworthiness for example, complaining about similar problems.
  - However, if the reviews are too similar, the readers suspect review fraud and deem the review as untrustworthy.
  - Great reviews posted right after negative ones give the readers the idea that the great reviews are a sort of damage control posted by the establishment itself, lowering the perceived trustworthiness.
  - Consensus with other reviews makes a review seem more legitimate (Munzel, 2016), meaning for example that posting a one-star review to a product with many great reviews undermines the reviews credibility.
- The medium in which the reviews are hosted on has great effect on the perceived trustworthiness.
  - Readers are skeptical towards corporate websites as they often seem whitewashed and overly positive.
  - Readers also cross-check reviews between different review sites.

In agreeance with Mathews Hunt, (2015), consumers see specific information posted by other consumers as the most trustworthy and useful as opposed to manager-generated content (Sparks et al., 2013). This is due to the information not originating from the reviewed company, making it highly credible in the consumers' eyes (Bickart & Schindler, 2001).

2.2 Falsified online reviews

With the online consumer reviews having such a huge effect on the consumers' purchase intention and companies' reputation, it is clear why false reviews can be a lucrative business.

A side-effect of increasing online review popularity, the false reviews are reviews that generally aim to boost or slander its target. The idea behind this obviously being to make a service or a product seem better or worse than it actually is, which makes the false reviews generally have a very high or very low rating score, compared to the truthful ones (Luca & Zervas, 2016). However, Martens and Maalej, (2019) state that some false review providers specifically ask the reviewer to not post a five-star rating in order to seem less suspicious.

Ranging from paid 5-star reviews on app-stores and Facebook groups to astroturfing campaigns ordered by multibillion dollar companies and everything in between, falsified reviews and opinion come in all sorts and sizes. The practice of purchasing reviews is generally seen as unethical and dishonorable and in the grey area in terms of legality. Despite this, the business regarding false and paid reviews is booming, whether it is an app developer paying a user for a positive comment on their Appstore page or a large-scale crowdturfing operation ordered by a global company.

While not all untruthful or spam reviews are written with malice in mind, they still can be expected to affect the companies' public image if left untended. Jindal and Liu, (2008) categorize these spam reviews into three types: untruthful opinions, reviews on brands only and non-reviews.

- **Untruthful opinions:** Deliberately misleading with the intention to boost or defame a target
- **Reviews on brands only:** Reviews are targeted at the brand instead of the product or service, thus being often biased.
- **Non-reviews:** This category contains ads, questions, answers, and other random texts.

Only the untruthful opinions are purposely falsified reviews of the three, the other types of reviews are also mixed in with the real reviews and affecting the overall rating thus being able to affect the customers' purchase behaviors.

While you generally can trust the bigger review platforms such as google and yelp, there are still malicious actors aiming to boost or negatively affect a review page of a business or a product. Streitfeld, (2012) estimates that a third of all online reviews are fake. Also, according to Luca and Zervas, (2016), 16% of all Yelp restaurant reviews are filtered, meaning they are suspected of being fake. Also, yelp itself has admitted that 25% of the reviews on the platform are atleast somewhat suspicious (Munzel, 2016). Considering false positive reviews are more popular than negative (Lappas et al., 2016), the baseline for review ratings on many platforms are artificially higher than if the reviews were 100% truthful (Byers, Mitzenmacher & Zervas, 2012).

While the exact percentages are impossible to know, it is obvious that false reviews exist in the online review platforms and communities. However, the estimated percentages may vary depending on the platform and detection methods used for example.

Also, it is worth noting that the numbers come from the detected or suspicious reviews while the "better" fake reviews go on undetected. However, these estimated values may also contain false positives, for whatever reason.

2.2.1 The effects of false reviews

Considering online reviews in general are very influential and trustworthy in the eyes of the consumer, the effects from falsified reviews can be considerable. While companies such as TripAdvisor and Yelp are hard at work trying to remove and prevent false reviews on their platforms, the "best" individual false reviews are the ones that have not been detected as is the case with most internet scams and frauds. The effects from these undetected false reviews can be difficult to determine as they are indistinguishable from legitimate reviews and thus can be considered to have the same effect as a legitimate review.

In a study of false reviews on Facebook, He, Hollenbeck and Proserpio (2021) discover that after a company ceases to purchase false reviews, their average rating rapidly drops and the number of one-star reviews increases dramatically. This was particularly true for young products. This may indicate that the buyers felt deceived, and that the product did not match the manipulated rating. This in turn creates an especially negative experience for the customer which causes them to be more likely to leave a bad review (Hu, Pavlou & Zhang, 2006).

A consensus among the academic literature is that false reviews undermine the value of online reviews and the reputation system itself. Being aware that a company has employed false reviews in order to boost their ratings, have been proven to lower consumers trustworthiness towards the reviews (Jin Ma & Lee, 2014) and damage the reputation of the involved companies (Ahmad & Sun, 2018). This in turn can cause problems for a company as trustworthiness strongly mediates purchase intention and future interactions with a service provider (Munzel, 2016). However, knowing that a company has received negative false reviews with the purpose to damage its reputation, increases the purchase intention as the false negative reviews are disregarded in the decision-making process (Jin Ma & Lee, 2014; Filieri, 2016).

A consumer being aware of false reviews on a certain has medium causes the false reviews to have an opposite effect on the consumer as to what was intended by the writer.

Even a review platform is not immune to the effects of false reviews, as in 2014 TripAdvisor was fined half a million euros in Italy for misleading reviews while claiming that the reviews on their site were legitimate and trustworthy (Masoni, 2014).

2.2.2 Characteristics of false reviews

In order to spread misinformation, the falsified online reviews have to be believable in the eyes of the reader, as the reviews that are found to be falsified are discounted from the decision-making process (Jin Ma & Lee, 2014; Filieri, 2016). The success of a false review is closely tied to the trust it invokes as the more trustworthy a review seems, the bigger weight it has on the decision-making process (Filieri, 2016). The factors affecting a reviews trustworthiness were covered in a previous chapter.

The prevalence of positive false reviews being higher than their negative counterpart (Lappas et al., 2016) is also reflected in the words used in the reviews, as the false reviews tend to *"…include more positives adjectives and less negative words related to software engineering such as "fix" or "crash""* (Martens & Maalej, 2019). While artificial intelligence exists to determine whether a review is fraudulent or not, from the perspective of the reader, distinguishing between real and fake reviews may be difficult (Kim, Kang, Shin & Myaeng, 2021). In fact, in an experiment by Ott, Choi, Cardie and Hancock, (2011) the highest scoring human was only able to detect 61% of the false reviews, even when the choices of words used in the reviews were different between the real and fake ones.

However, there are cues not related to the reviewer or the review itself that may indicate the presence of false reviews, such as a bimodal review graph or a sudden burst of very good or bad reviews. Many reviews written in a short amount of time can be an indication of false review campaign, as the false reviewers tend to work together according to Li et al., (2017). While a sudden burst of reviews has been found to be suspicious in the eyes of the reader (Filieri, 2016), the reader may not notice such an event without taking a closer look at the reviews and their post times.

The ratio between negative and positive false reviews may depend on the product/service category. According to Luca and Zervas, (2016) restaurants are more likely to get false negative reviews as the competition increases, while Martens and Maalej, (2019) who studied false reviews on the Apple Appstore concluded that the amount of false negative reviews was extremely small (1- and 2-star false reviews making up only 1.6% of all false reviews). However, in the case of appstores, it may be much harder to figure out the direct competitors within the abundance of apps compared to a few local restaurants. All in all, as false reviews are difficult for humans to detect reliably despite the apparent confidence in doing so.

2.2.3 Bimodal reviews

While detecting false reviews has been proven to be very difficult for humans (Kim et al., 2021), apart from obvious ones, a common symptom caused by false reviews is a review graph with two peaks, as false reviews tend to favor more extreme ratings in order to boost or slander the target (Luca & Zervas, 2016). This phenomenon is called bimodal reviews, a tangible effect that can be seen warping the different reviews graphs and scores.

When a company aims to boost its own reputation through false reviews, especially if the starting reputation is poor (which is a more likely starting point than a higher rating according to Luca and Zervas, (2016) and Chen, Li, Chen and Geng, (2019)) the result is a distinctively shaped figure with two peaks, one from the legitimate poor reviews and one from the falsified positive reviews for example.

The average rating-score may not show the true nature and quality of the product for multiple reasons. The ratings might have been spoofed one way or another, or the product/service is highly polarizing, such as books, videos and DVDs as was the case in the study by Hu, et al., (2006). The innovators and early users' opinion can also vary greatly from the opinions of the majority, warping the "real" rating (Li & Hitt 2008). Another factor related to this is that buyers tend to leave reviews after a particularly bad or good experience, a phenomenon Hu et al., (2006) titled "brag-or-moan". These factors, among others can make the average rating-score too high or too low for what the product/service is.

To give insight to the rating of a product/service a rating distribution is typically provided. This is usually depicted as a bar chart ranging from 1 (worst rating) to 5 (best rating). While most legitimate products/services have a clear peak in the rating distribution called a unimodal rating (Luca & Zervas, 2016), a bimodal rating tends to raise suspicion. Why do some people seem to be very content with the product, while others completely hated it? Of course, in some cases the bimodal distribution of reviews can be legitimate due to the factors stated earlier. However, as stated by Luca and Zervas, (2016) the effect of false reviews regarding bimodal reviews is substantial and cannot be ignored. They argued that false reviews may account for a substantial number of extreme reviews (1 or 5). Studying data from filtered (removed on suspicion of being fake) yelp reviews, they found out that almost 20% of filtered reviews were 1-star and almost 40 % of them were 5-star reviews while accounting for under 10% and 30% of the legitimate reviews respectively.

2.2.4   Crowdsourcing the reviews

In crowdsourcing, a task is distributed to a large audience to take advantage of the masses (Lee et al., 2015). While no clear definition of crowdsourcing exists, in its simplest definition it is to outsource an activity to the crowd (Estellés-Arolas & González-Ladrón-de-Guevara, 2012). Typically crowdsourcing is used to complete tasks that are easier for humans to complete than for computers (Tong, Zhou, Zeng, Chen & Shahabi, 2019). Such tasks include object identification in videos or images and audio transcribing for example (Ipeirotis, 2010a). Crowdsourcing itself is a legitimate business model and only becomes malicious when the task itself is, as is the case with writing false reviews

The Amazon Mechanical Turk, (from here on AMT) is an example of a platform that provides crowdsourcing services to their customers, who are referred to as "requesters". These requesters, such as companies, send their tasks, referred to as human intelligence tasks (HIT) to the remote workers through the AMT platform. Another similar service provider is Fiverr.com. The website does not

advertise itself as a crowdsourcing service as it consists of freelancers offering a multitude of services to anyone willing to pay as reported by USA today (Weise, 2015). Main difference to the AMT being that Fiverr has individuals offering a certain type of work, while AMT offers workforce for the requesters.

Offering incentives for writing reviews is also a way of crowdsourcing reviews, which is against many review platforms' guidelines, for example Amazon prohibits any kind of compensation in exchange to writing reviews in their community guidelines ("Amazon.com Help: Community Guidelines", 2021). However, some exceptions are made, for example book authors may offer a free or discounted copy to a reader as long as they do not specifically request a review or try to affect it. In any such case, Amazon requires the financial connection to be disclosed. In a survey by Brightlocal (Murphy, 2020), 12% of respondents had been offered a discount as an incentive for reviews, money for 9% and gifts/free services for 8% respectively.

2.2.5 Astroturfing

Astroturfing is the act of using fake entities to generate desired attention to the target, making one's opinion look like it is coming from multiple different legitimate users (Zhang, Carpenter & Ko, 2013). For example, using bots to inflate a restaurant's reviews to make it seem better than it is, or vice versa, orchestrating a campaign to give poor reviews to the competing firms. The analogy for the name astroturfing, is derived from the artificial grass making company AstroTurf, symbolizing the use of fake grass (fake users) to represent real grass (real users).

Astroturfing false reviews doesn't always occur through platforms however, nor does it only exist in smaller companies. In a rather recent incident, in 2013, Samsung paid a notable number of writers to criticize the competing HTC phones and praise its own products. The Korean tech company was found to have enlisted two local marketing companies to write the false reviews on a Taiwanese forum as reported by the tech news website The Verge (Souppouris, 2021), among others. Samsung was issued a fine of $340,000 by the Taiwan's Fair-Trade Commission (FTC). The article also mentions that the FTC found Samsung guilty of "*disinfection of negative news about Samsung products*", which likely refers to calling poor reviews about Samsung products liars or fake and posting positive reviews of Samsung products.

2.2.6 Crowdturfing

According to Lappas et al., (2016), false reviews are largely a product of professional review companies, who orchestrate the campaigns to spread the falsified reviews for a fee. As writing credible false reviews is difficult for computers (Lee et al., 2015), humans must do the work instead, which leads to the phenomenon of crowdturfing. Crowdturfing is a form of astroturfing, in which real people, instead of bots, are paid to spread a desired opinion about the target entity (Rinta-kahila & Soliman, 2017). It is to be noted that crowdturfing does not consist solely

of false reviewing but includes other actions, such as commenting or liking a post which also have the effect of spreading one's opinion.

Crowdsourcing, and thereby crowdturfing has become popular due to its ability to complete tasks that are difficult for computers to accomplish (Lee et al., 2015). A typical crowdturfing campaign is conducted by an agent, or the "service provider", such as a crowdsourcing platform, who enlists workers, ie. people to write the false reviews for the customer (Wang et al., (2012).

The growing popularity of crowdturfing can be observed by comparing the results of Wang et al., 2012 and Héder, 2018, both of which studied the same crowdsourcing website six years apart. In these six years, the number of listed campaigns on microworkers.com has risen from 267 a month (Wang et al., 2012) to an average of 619 a month (Héder, 2018), while the percentage of estimated crowdturfing tasks remained the same at around 90%.

The payment for these menial tasks is typically around ten cents (in US dollars) with a payment of more than $1 being very rare (around 0.02% of all listings) according to Héder, (2018), who studied crowdsourcing campaigns posted on microworkers.com website.

While crowdsourcing sites such as Amazon mechanical turk and Fiverr are legitimate businesses, a large amount of crowdturfing tasks still exist within them. According to a blog by Ipeirotis, (2010b) the amount of spam HITs in AMT in 2010 was found to be 40%. The spam included false review requests among the other spam requests. However, two years later, a study by Wang et al., (2012) found the percentage of crowdturfing on AMT to be only 12%.

While the amazon mechanical turk was an outlier with a relatively low percentage, other crowdsourcing websites had a larger percentage of crowdturfing listings, with as high as 90% in Zhubajie and Sandaha, the two largest Chinese crowdsourcing sites (Wang et al., 2012). The smaller US based crowdsourcing sites studied by (Wang et al., 2012) also had a relatively high percentage of crowdturfing, ranging between 70 and 95 percent.

2.2.7 The market for false reviews

With writing false reviews being such a lucrative and common business, there obviously must be a customer. That customer can be a company looking to boost its reputation or to slander its competition. Regardless, there are different factors and attributes related to the company that affect its willingness to collude with the false reviews. The following sub-chapters look at what those different factors and attributes are and also what are the factors and attributes that deter companies from colluding with the false reviews.

The top three categories for which consumers use online reviews for are computers, clothing and hotels according to Jin Ma & Lee, (2014), however a more recent survey by Brightlocal (Murphy, 2020) lists the top five as restaurants, hotels, medical, automotive, and clothing stores. The two are seemingly similar with the likely explanation for the popularity for the medical services being the prevalent COVID-19 pandemic. While the top categories for reading reviews are quite clear, the false reviews do not seem to be tied to any specific product

category, according to a study by He et al., (2021), who studied the market for false Amazon reviews. Despite this, false reviews were most prevalent in beauty & personal care and household related categories but overall, very widespread among the studied categories. However, in the case of Apples Appstore, an overwhelming amount of the false reviews seem to reside within the games-category (Martens & Maalej, 2019). The games-category represents around 23% of the apps on the apple Appstore but has just under half of all legitimate reviews at 49.95%. In their study Martens and Maalej, (2019) found that of all apps containing false reviews, around 53% fell under the category of games while containing just under half of the false reviews detected in the study at 47.57%.

2.2.8   Sellers of false reviews

False reviews are typically a product of professional review firms, who orchestrate the campaign for the customer (Lappas et al., 2016). These services are sold on for example crowdsourcing websites such as amazon mechanical turk alongside the legitimate listings. The requests for false reviews are then filled by individuals who get paid by the comment, in a phenomenon called crowdtrufing, which was introduced in a previous chapter.

The false reviews are also sold on third party websites not related to the false reviews themselves, such as in large private groups on Facebook (He et al., 2020). These Facebook groups contained on average 16000 members and 568 daily requests for false reviews. Martens and Maalej, (2019) found that most (79%) of the false review providers sell their services in exchange for payments, while the rest of the providers (21%) operated on the principle of exchanging reviews with each other.

A study made by Lee, Webb and Ge, (2015) lists different social media sites that crowdturfing seller's target. The top three being Facebook at 50%, Twitter at 26.6%, and Youtube at 12.5%. However, it is to be noted that the study was conducted in 2014 and platforms in the lower ranks, such as Instagram has increased its popularity massively, while services such as google+ have since ceased to exist. Also, the study explicitly studied social media sites, not regarding review platforms in their study. Nevertheless, as the main idea behind crowdturfing tasks targeting social media sites is to artificially generate likes, comments, and followers (Lee et al., 2015), which in turn gives a more positive outlook of the target, they are closely related to the false reviews.

As buying false reviews and other online impressions is typically seen as a shady practice, sellers of these services often have to reassure the potential buyer of their service. Rinta-kahila and Soliman, (2018) found that these sellers use a combination of three different persuasive strategies: educational, bragging, and reassuring messages. These strategies are not specific to these "shady" services however, as legitimate businesses also utilize them in their communications according to Rinta-kahila and Soliman, (2018).

2.2.9   Buyers of false reviews

For sellers of false reviews to exist, there must also be buyers. Chen et al., (2019) found out that the following characteristics lessen a company's tendency to collude with false reviews: reliance on repeat purchase, stiff penalties for fake reviews, low priced low-quality products when competing, high quality products are much more expensive, and if the seller already has a good reputation. They also found the following factors that increase a seller's tendency to collude with false reviews: low product score, high priced product, and fewer reviews on a product. The latter was also noticed by Martens & Maalej, (2019), who found that Apps that have false reviews on the Apple Appstore typically have between 2 and 9 reviews in total.

Luca & Zervas, (2016) also found out several factors affecting a companies' likelihood of posting false reviews. For example, companies with weaker reputation were more likely to commit review fraud, while established chains were far less likely to do so. Also, direct competition was observed to increase negative reviews aimed at the competitors. Despite all the factors that would suggest that review manipulation is focused on smaller and less popular companies, it is worth noting that even the largest companies in the world may not shy away from these dirty tactics considering the Samsung astroturfing scandal in 2013 (Souppouris, 2021).

To reiterate, factors that were found to make false reviews more lucrative were: direct competition, weak product reputation, weak company reputation, low number of reviews on product, and high price product. On the other hand, several factors that made false reviews less lucrative option were: reliance on repeat purchase, stiff penalties, good reputation, much lower price compared to competition, new product, and chain affiliation.

He et al., (2021) results also point, that new products rarely purchase false reviews in order to boost their reputation with the average product age being 229 days and median age being 156 days for products with false reviews. However, it is debatable what can be considered new in which circumstances. For example, a restaurant that has been up for around five months is would likely be considered new by most.

Considering that the aim of the false reviews is either to boost a company's review rating by writing positive reviews, or to lower it by posting negative reviews, the underlying motivation is monetary gain and maximizing profits (Wu, Ngai, Wu & Wu, 2020), which is also affected by competition to an extent.

As the crowdturfing actions are done by ordinary people, it is important to understand the rationalizations that one makes when participating in this type of cyber deception. Through a vignette study, Kauppila and Soliman, (2022) present five main techniques that are used to justify the act of cyber deception. While the first three techniques are based on the five neutralization techniques by Sykes and Matza, (1957), appeal to professionalism and appeal to normative fluidity were new categories found in this study. However, the 'condemnation of the

condemners' and 'denial of the victim' techniques by Sykes and Matza, (1957) were found to not be acceptable justifications in this case. The techniques, their explanations, and whether they are an acceptable justification in the case of false reviews is presented in table 1.

A common theme in the interviewees' answers seems to be the belittling and ignorance of the effects of cyber deception and false reviews, thus raising awareness about said consequences could be an effective way to combat the phenomenon from an ethical perspective. While the study by Kauppila and Soliman, (2022) looks at the false reviews' writer's perspective, some of the factors could be seen to concern the buyer side as well. For example, denial of injury when the case is not about smear campaigns for obvious reasons.

Considering that for the individual to participate in false review crowdturfing campaigns, they first must justify their actions, for example through aforementioned rationalizations. This act of rationalizing could be seen as the first hurdle to overcome when looking at the overall roadmap of a false review and the security action cycle, which will be covered next.

Table 1 Justification techniques for false reviewers

Technique	Explanation	Acceptable justification
Appeal to higher loyalties (Sykes & Matza, 1957).	For example, monetary gain was seen as an acceptable reason to participate, while helping a friend was seen as a less legitimate reason. In this case, the writer of the false reviews was seen as providing for themselves.	Yes
Denial of responsibility (Sykes & Matza, 1957).	Not being aware about rules related to cyber deception and different loopholes were generally seen as passable rationalizations, however some interviewees felt that the rules should be common sense and the rationalization was just an excuse.	Yes
Denial of injury (Sykes & Matza, 1957).	Many interviewees saw the act of cyber deception as nigh victimless.	Yes
Appeal to professionalism (Kauppila & Soliman, 2022).	By using this technique, a person practicing cyber deception passes the blame to a higher entity, such as an employer as "they are just doing their job".	Yes
Appeal to normative fluidity (Kauppila & Soliman, 2022).	This rationalization refers to the ever-changing norms. While this type of action might have been illegal or	Yes

	frowned upon earlier, it might be more acceptable and prevalent in the future.	
Denial of the victim (Sykes & Matza, 1957).	The act is framed in a way that the victim deserved it, such as vengeance or as a 'Robin Hood-like' action.	No
Condemnation of the condemners (Sykes & Matza, 1957).	The attention is shifted towards the condemners, framing them as hypocrites and spiteful towards the actor for example.	No

3   SECURITY ACTION CYCLE AND SYSTEMS RISK

A way to study companies false review preparedness is to view it through the security action cycle perspective. Originally based on Nance and Straub (1988), Straub and Welke, (1998), proposed the security action cycle for managing systems risk, which they define as: "*The likelihood that the firm's information systems are insufficiently protected against certain kinds of damage or loss (…).*". The security action cycle consists of four stages/phases: deterrence, prevention, detection, and remedies. Each of these stages are explained in the following chapters.

Each stage contributes to deterring future computer misuse through a deterrence feedback loop, which makes the potential violators aware of the consequences. The deterrence feedback loop feeds the deterrence stage through the actions from the other three phases. The prevention, detection, and remedies stages contribute to the deterrence stage by making a potential offender aware of the consequences and potentially thus deterring them from committing the unwanted action. (Straub & Welke, 1998). For example, the knowledge that a company might file a lawsuit in regard to false reviews or that the false reviews will be swiftly dealt with are just some examples of the ways that the different phases provide to the deterrence feedback loop and ultimately serve the deterrence stage.

Each stage also has a specific objective on top of contributing to the deterrence feedback. The deterrence and prevention stages aim to maximize the deterred and prevented abuse respectively, while the detection and remedies stages aim to minimize undetected and unpunished abuse respectively.

This chapter views and explains the security action cycle by Straub & Welke, (1998) and how it can be applied to be used against false reviews.

A viewpoint not included in the security action cycle is how the writer of false reviews justifies their actions to themselves. As Kauppila and Soliman, (2022) found, there are several different ways for the writer to justify their actions through rationalizations, which could be seen as the precursor for the actions related to the security action cycle. These rationalization strategies were presented previously in table 1, which also included the neutralization techniques by Sykes and Matza, (1957). The rationalization stage would weed out the possible

crowdturfers and false reviewers from even engaging in the fraudulent act if they could not find a justification for their actions. The figure 1 depicts the security action cycle as developed by Straub and Welke, (1998) with the added rationalization pre-stage, which is based on Kauppila and Soliman, (2022) and Willison and Warkentin (2013). The figure also clarifies at which point the false review becomes public.
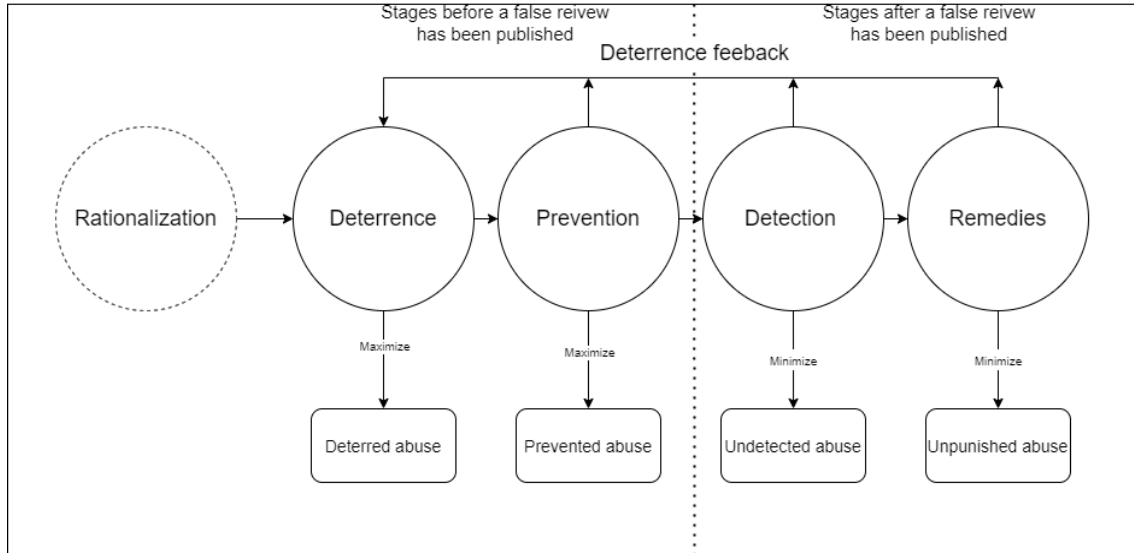
Figure 1 The security action cycle by (Straub & Welke, 1998) with added rationalization phase based on (Kauppila & Soliman, 2022).

3.1 Deterrence

The deterrence phase consists of passive factors, such as guidelines and rules that users are expected to follow. The deterrents themselves do not exactly prevent users from performing unwanted actions but serve as warnings for potential offenders. Deterrence has a strong theoretical basis, with mentions to as early as in the 1800's (Anderson, Harris & Miller, 1983). An example of a more recent study, Lembcke et al., (2019) look at deterrents affecting information security compliance and list three factors: perceived sanction severity, certainty and celerity based on the reviewed literature. A closer inspection of deterrence theory to inspect the aforementioned factors would be beyond the scope of this study, however.

A deterrence theory explains what negatively perceived factors deter an individual from doing a certain action, which can further be divided in to specific and general deterrence depending on if the punishment has been personally experienced (specific) or just observed (general) (Siponen, Soliman & Vance, 2022). The case with the security action cycle is that the different stages can be considered to provide for the deterrence stage through general deterrence as they serve as "examples" of what will happen or has happened to others.

The main objective of the deterrence stage is to maximize the deterred abuse with the help of the other stages which provide for the feedback loop (Straub & Welke, 1998). The purpose of the deterrents is to convey that the company is serious about policing computer misuse. In the realm of false reviews and review abuse, this stage consists of the warnings that: the false reviews are not allowed, they will be deleted and that the offenders will be punished for example. In short, this stage can be seen as the "beware of the dog sign".

When discussing about deterrence relating to false reviews, it is important to look at how the platforms deter false reviewers. The rest of this chapter looks at the punishments issued by the platforms in which the reviews exist in. By examining some of the most popular review platforms such as TripAdvisor and Yelp and other platforms (from here on collectively referred to as review platforms) containing large amounts of reviews such as google maps and the Apple App store, a plentiful understanding of the different punishment methods and sanctions can be acquired. It is to be noted that these punitive actions can also be seen to belong in the remedies phase but due to the deterrence feedback loop, they are introduced in relation to this phase.

False reviews are generally prohibited on review platforms in order to provide the consumers with unbiased opinions about the companies and their services as well as to preserve the reputation of the review platforms themselves ("App Store Review Guidelines - Apple Developer", 2021); ("What does Tripadvisor do about unfair reviews?" (n.d.)); ("What is Yelp's review solicitation penalty? | Support Center | Yelp", 2021)). To combat review fraud, the review platforms have adopted deterrents to discourage the companies from engaging in this fraudulent behavior. However, some platforms may have a requirement for posting reviews, such as proof of purchase or a verified buyer badge. This is called high posting cost and it is discussed in more depth in the prevention chapter under security action cycle and systems risk.

The deterrents themselves do not prevent the false reviews as they consist of guidelines, rules and the like according to the security action cycle by Straub and Welke (1998). The banners issued on offending companies pages also serve as general deterrence, meaning a punishment has been observed to have been issued on someone else (Siponen, Soliman & Vance, 2022).

The punishments vary from platform to platform but in general a similar toolset is in use on every platform to deal with false reviewers and their collaborators. Punishments include for example, banners on the company's pages on corresponding platforms, fines, account deactivations, and even jail time. Table 2 lists punishment and disciplinary actions used by different platforms.

As a response to the growing problem of false reviews, TripAdvisor has cracked down on false review sellers affecting their platform. Reportedly a first of its kind, in 2018 the owner of an Italian company PromoSalento, was sentenced to jail time by the Criminal Court of Lecce as they were found guilty of committing a fraud by using fake identity by selling false reviews (Kempf, 2018). This was in part due to TripAdvisors efforts in shutting down the company.

TripAdvisor has implemented several penalties for companies found guilty of colluding with false reviews, such as a banner on their page, ranking penalties and ineligibility for being highlighted as a top choice (TripAdvisor, 2021). The banner placed on an offending company's page reads:

> *Message from TripAdvisor: TripAdvisor has reasonable cause to believe that individuals or entities associated with or having interest in this property may have interfered with traveler reviews and/or the popularity index for this property. We make our best efforts to identify suspicious content and are always working to improve the processes we use to assess traveler reviews (("Review Fraud Consequences," (n.d.))).*

The review platform Yelp issues similar types of penalties to the offending parties as TripAdvisor. Search result ranking penalties for companies found to manipulate their reviews, making them appear later in the search results and in more severe cases apply a banner to the offending companies Yelp page for three months ("What is Yelp's review solicitation penalty? | Support Center | Yelp", 2021); (Luca & Zervas, 2016).

If a developer is found to manipulate reviews or otherwise inflate their ratings on Apples Appstore, they risk getting expelled from the Apple Developer Program and having their app removed from the store ("App Store Review Guidelines - Apple Developer", 2021). However, according to a study by Martens and Maalej, (2019), only 7% of the apps that contained false reviews according to the study were removed from the Apple Appstore in the timespan of three months. Apple does not elaborate on the different punishment methods beyond the extreme ones mentioned earlier.

Amazon on the other hand elaborates on the different punishments to a slightly greater extent on their policy page ("Amazon.com Help: Anti-Manipulation Policy for Customer Reviews", 2021). Amazon lists punishments for companies trying to manipulate their ratings as: account suspension/termination, review removal and product delistment. Amazon also mentions that if a specific amazon account has been found to manipulate reviews, their "–*remittances and payments may be withheld or forfeited.*". Amazon also actively pursues lawsuits against the review manipulators, as false reviews may be illegal depending on the state and federal laws in the US.

Google reports about its anti-false review battle in it's the Keyword blog, written by Principal Software Engineer Dan Pritchett (Pritchett, 2021). According to the blog, Google disables the accounts of policy-violating and suspicious users. In 2020 alone, 300000 business accounts were disabled due to being reported for policy violations. Most of these violations belong to one of two different categories: content fraudsters and content vandalism. These two categories correspond to the untruthful opinions and non-reviews respectively as coined by Jindal and Liu, (2008). While stiff penalties were found out to be an effective countermeasure by Chen et al., (2019), Lappas et al., (2016) note that these types of penalties can cause companies to frame their competitors in order to damage their reputation and visibility.

Table 2 Review fraud punishments by platform

Company/platform	Punishment method
TripAdvisor.com	Banner, ranking penalty, ineligibility for highlights
Yelp.com	Banner, search ranking penalty
Apple Appstore	Expulsion from dev program, app removal from store
Google maps	Business account deactivation
Amazon.com	Business account suspension/termination, product delistment, comment removal. For individual accounts withholding or forfeiture of funds. Lawsuits towards the false reviewers

3.2   Prevention

This stage consists of the factors that aim to prevent abuse in cases where the offender has chosen to ignore the warnings. The objective of this stage is to maximize the amount of prevented abuse, while contributing to the deterrence feedback loop by making potential offenders aware of the preventive measures (Straub & Welke, 1998). In the case of false reviews, such factors can be different requirements on websites for posting reviews. For example, Ott, Cardie and Hancock, (2012) found that having a high posting cost, meaning that the user must for example purchase the product before they are able to review it, lowers the amount of review spam on the platform. Another example of a posting cost is the popular online forum Reddit.com, which requires users to have enough "karma" to post in some communities. In Reddit, "karma" is earned by posting and commenting on other users' posts. This is to prevent spammers from creating new accounts and swarming the website according to Reddit itself ("Why am I being told, "you're doing that too much..."?" (n.d.)), which is in line with Ott, et al., (2012) according to whom, users with low amounts (1 or 2) of reviews are more likely to post spam reviews than those with a higher number of reviews posted. While Reddit is not a review website, the effect of eWOM can be expected to be

present, as is the case with all online forums. It is to be noted that, while both examples have a posting cost, having to purchase a product before reviewing it has a much higher posting cost than acquiring enough karma in the case of Reddit.

While a high posting cost such as verification of purchase seems like a way to cut down false reviews Lappas et al., (2016) argue that it may bring the following negative effects. Firstly, a company wanting to boost their own reviews buys the product from themselves, making the cost return to the company. Secondly a motivated false reviewer can still purchase the product, especially in low-cost market. Thirdly the buyer and sales channel affect who can post the review, for example only a single family member can comment on a hotel room because they were the one to make the reservation. Also, if the reservation/product was purchased from a third party, they might not receive the verified buyer status. The verified buyer status may be the most effective versus astroturfing campaigns as individual workers likely do not want to purchase the product just to leave a single review. (Lappas et al., 2016).

Also, automatic false review detection systems (not to be confused with the detection stage) can be considered as preventive mechanisms, as an AI can detect some false reviews through text-based detection mechanisms and prevent them from appearing on the website. TripAdvisor for example has all its reviews checked by smart technology that blocks, allows or flags the review for further inspection by a human moderator in unclear situations ("What does Tripadvisor do about unfair reviews?" (n.d.)).

Amazon has taken more drastic measures to prevent false reviews from being written by filing lawsuits against over 1000 people who were offering false reviews on the crowdsourcing platform Fiverr.com (Weise, 2015). While this could be seen to belong in the deterrence stage through general deterrence by deterring someone from even trying, it can also be seen to belong in to the prevention stage through specific deterrence, which prevents someone from posting false reviews repeatedly.

Lappas et al., (2016) identify two different response strategies for companies dealing with false reviews: enhancement and confrontation, with the enhancement strategies containing three sub-categories. The study in question was conducted on the domain of hotels but it could be applied to different domains as well. This section summarizes the enhancement response strategy proposed by Lappas et al., (2016) and based on which figure 2 is created. As the confrontation strategies deal with suspected false reviews that have gotten through whatever safety systems were in place, it will be addressed in the remedies chapter under the security action cycle.

As the name suggests, an enhancement strategy focuses on making the business look better and have more visibility than its competitors in the eyes of the potential customer. This strategy can also be applied pre-emptively to improve the company's standing instead of reactively as a remedy for false reviews. This strategy can be implemented through three different sub-categories: rising

through the ranks, covering more features, and improving the quality of existing features

The first sub-category of the enhancement strategy is called rising through the ranks, which is based on the company having a better ranking than its competitors on any relevant ranking site for example. This enhancement can be done for example through a marketing campaign which in turn draw in more reviews according to Byers et al., (2012). However, the reviews must be positive to affect the company's ranking positively. Lappas et al., (2016) also note that this strategy can be implemented through false positive reviews as well as through false negative reviews on the competitors who have a better ranking. Fighting false reviews with false reviews can potentially develop into a vicious circle of false reviews between competing companies and should thus be discounted from being a possibility by the employees managing company's reviews.

The second sub-category in the enhancement strategy is to cover more features, which in turn translates in to covering the needs of more customers. However new features are often costly and not very easily obtainable, depending on the type of business and domain. It is also important to consider what features are wanted by the potential customers to not waste resources on a feature no one appreciates. A company can choose between imitating an existing feature from a competitor and differentiating by innovating a completely new feature when deciding on a new feature (Narasimhan & Turut, 2013).

The third and final sub-category in the enhancement strategy is to simply improve the existing features and services to gain better reviews and prevent the poor ones from being written. As is the case with adding new features, the company aiming to improve its existing services should prioritize the features that their customers value the most.
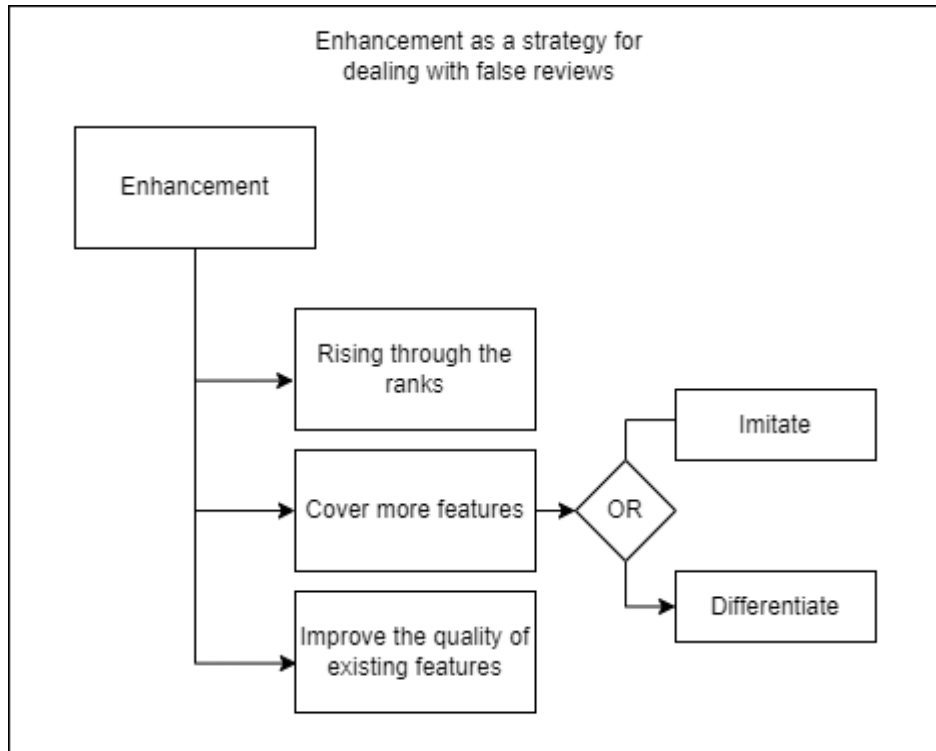
Figure 2 Enhancement strategy based on (Lappas et al., 2016).

3.3 Detection

If the offender chooses to ignore the deterrents and manages to bypass the preventive mechanisms, the next line of defense is detection. At this stage, some damage has already been done, for example a false review has been posted because the website did not recognize the user as a bot or there were no preventive measures in the first place. The purpose of this stage is to minimize the undetected abuse while contributing to the deterrence feedback loop by demonstrating that the offenders will be detected and dealt with accordingly (Straub & Welke, 1998).

Currently the detection of false reviews focuses on two aspects: text-based and reviewer-based perspectives (Luo, 2020). Zhang, (2020) concludes that using user behavior data to catch fake reviews is better than text-based data through machine learning. It was also found to be better due to the fact that processing text data is very slow compared to processing user activity data.

Reviewer based perspective: As the name suggests, this method looks at the entity posting the reviews. According to Ott et al., (2012), reviewers with one or two posted reviews tend to be more likely spam than reviewers with more postings. Also, the probability of spam with one review reviewer was higher than with two review reviewers. Another example of reviewer-based detection method was made by Lee et al., (2015), who have made classifications to detect

the crowdturfers on Twitter with 99.29% accuracy. However, their detection method also included the users' text analysis.

Text based perspective: The text-based detection methods are based on linguistics, statistics, and computer science (Luo, 2020). They focus on detecting the fake reviews based on the review length, rating (Mukherjee, Venkataraman, Liu and Glance, 2013) and linguistic factors such as readability and writing style (Banerjee & Chua, 2014).

While larger review platforms and companies can have different types of AI filtering, flagging, and removing the false reviews, the smaller firms may not have the resources. Also having sophisticated AI filtering reviews on a small local store's website may feel like an overkill. The language barrier could prove an issue, as this study is conducted in Finland, most of the relevant reviews should be expected to be in Finnish. A language processing AI might have issues with this, not only because of a different language but due to the different dialects which are also prevalent in the written language. However according to Googles, The Keyword blog (Pritchett, 2021), Google employs thousands of human analysts to understand reviews written in local slang/dialect but it is unclear if this only applies to the English language.

A common detection method among the different review platforms, that applies only after the review has been posted, is the ability for users to report suspicious reviews and users. For example, TripAdvisor removes the reported reviews if they are found to violate their guidelines ("What does TripAdvisor do about unfair reviews?" (n.d.)).

Customers who are highly involved with a certain product purchase may start to detect patterns in the reviews according to Filieri, (2016). While not a sure sign of false reviews, a bunch of glowing reviews posted in a small timeframe atleast raises suspicion of manipulation, especially if they are posted after poor reviews.

The detection phase of the security action cycle should not be confused with the detection of false reviews, which can be considered to belong to both the prevention and detection phase, depending on when the detection of false reviews happens.

3.4 Remedies

The final stage of the security action cycle is remedies. This stage focuses on recovering from the incident and punishing the perpetrators. The objective of this phase is to minimize the unpunished abuse, while contributing to the deterrence feedback loop by demonstrating that offending parties will be punished for their actions (Straub & Welke, 1998).

In the case of false reviews, this stage consists of dealing with the false reviews after they have been noticed. Different review platforms have different punishments for the offenders but the most common actions such as removing the reviews, banning users from the platforms, and issuing banners on the

offending companies pages on the review website. Other more severe sanctions may include lawsuits as mentioned earlier and other legal action, as false reviews are considered illegal in many countries. Punishments and sanctions issued by different review platforms for review fraud were presented earlier in table 2.

Responding to the suspected false reviews is also a main point of the confrontational strategies by Lappas et al., (2016) who focus on dealing with the detected false review as a remedy. This strategy intends to deal with the false reviews by reporting or responding to the suspected false review to challenge its message. Tailored responses to negative reviews have been proven to improve the future opinions of the said review (Wang & Chaudhry, 2018), mitigating the damage done by the false review to a certain extent. Responding and challenging the false reviews should be the initial response to false reviews, as it may take time for the review platform to assess the situation and remove the false review after it has been reported.

Depending on the platform, the verified badge often given to the business accounts on third party websites has been associated with increased trust towards the writer (Kim et al., 2017), making the response have more weight. A company suspecting false reviews on their page on a third-party website should flag the review for further examination by the site's moderators for example.

Also, in the case of review platforms, actively removing false reviews and displaying commitment towards a truthful environment can act as a token of trust for the users of the platforms.

4 RESEARCH METHOD

This chapter describes the empirical research done in order to acquire the required data for the study. The empirical part of this thesis is conducted as qualitative research through semi-structured interviews. As the aim of this study is to find out what action's companies are taking to prevent false reviews from having an impact, the logical way for that is to inquire the individuals in charge of the companies' social media, reputation management, and other relevant online customer relations. In other words, the employees who often interact with customers online in a public setting or know about the company's operations regarding it.

4.1 Design

Interviews were chosen for this purpose, as they are the most suitable for extracting in-depth information from an individual (Fontana & Frey, 2000). Also, as the secondary research question looks at reasons why the companies would not collaborate with false reviews, the interviews allow for probing of ones more personal opinion on the subject.

A semi-structured interview method was chosen over structured and unstructured interviews, since the required data is based on the security action cycle by Straub and Welke, (1998). In this situation an open interview might not answer the required questions and a structured interview might not allow for the interviewee to tell in sufficient breadth about their situation. Having some questions to guide the interview, while allowing the interviewee to get more in depth with the subject is a main strength of the semi-structured interview (Myers & Newman, 2007), making it optimal choice for this type of study.

The security action cycle by Straub and Welke, (1998) depicts how future computer misuse is deterred by the actions of the different stages in the cycle. This can be applied into the realm of false reviews as the act of writing false reviews can be viewed as computer misuse due to its immorality and to some extent, illegality. The questions used to guide the interview are categorized based

on the security action cycle in an attempt to gain enough insight on all of the phases of the cycle. These questions are worded as unambiguously as possible in order to make them more understandable for the interviewee and to prevent misconceptions.

The sample for the study consists of 8 case companies. These companies reside in the service industry, with the interviewees consisting of restaurant and hotel representatives for example. The interviewed companies were chosen on the basis of not being a part of a chain franchise, as they are less likely to be affected by reviews due to rigorous marketing (Luca, 2011). Also, a simple pre-screening was done through different review and social media sites to ensure the interviewed companies had an online presence and some experience with reviews.

4.2   Forming the questions

The questions are categorized based on the security action cycle by Straub and Welke, (1998). Each of the four stages, deterrence, prevention, detection, and remedies contain relevant questions to figure out how the company in question utilizes said stages. In order to answer the second research question, the interviewee is asked about reasons not to collude with false review providers to find out what factors deter them from this practice. Also, some introductory questions about false reviews are asked, such as the interviewees experience with them and what review platforms etc. are in use by the company.

The questions used to guide the interviews can be seen in the appendix in both Finnish (appendix 1) and English (appendix 2). On top of dividing the questions by the phases of the security action cycle and research questions, some questions were asked based on imaginary situations and resources. Also, what responsibilities are transferred to the review platforms and what the interviewee thinks they are doing or should be doing regarding the false reviews are discussed.

The first two interviewees, Interviewee 1 and Interviewee 2 can be considered as sort of pilot interviews, as the questions were slightly modified based on their answers and hence some questions were not asked of them. As a result of the revision a few questions were added but nothing that would make a significant difference between the first two interviews and the rest.

4.3   The sample

The sample for the study consists of eight case companies whose representatives were interviewed. The interviewees were mid to top level management in their respective companies. The interviewees are from here on referred to as Interviewee (number) which is based on the order in which they were interviewed

Interviewee 1 being the first and Interviewee 8 being the final interviewee. The criteria for the interviewees were to work or have worked in a company that provides goods or services for the consumers in a position that involved reputation management such as reviews, social media or equivalent. Larger chains and franchises were excluded from this sample, as they are less susceptible to false reviews (Luca, 2011). The underlying idea is to discover the information about companies' false review preparedness and opinions through the people whose work is closely related to them. The initial target was to have ten participants, however the recruitment process turned out to be difficult, as many restaurants and bars for example were facing a severe staff shortage due to covid-19 among other reasons that emerged during recruitment.

The size of the company was also considered when choosing the companies. However, due to the low response rate from contacted companies and the information prior to the interviews about the companies' exact size was difficult to find, this factor ended up having less focus in the end. However, a rough divide can be made in terms of size. Four of the largest companies were Interviewee 2, Interviewee 4, Interviewee 5, and Interviewee 8. Two of the case companies, Interviewee 2 and Interviewee 5 were clearly the largest, with both having 50 or more employees in total. Interviewee 4 was slightly smaller with 25 to 50 employees and Interviewee 8 being in the 10 to 25 employee range. Rest of the case companies were notably smaller with Interviewee 3 and Interviewee 7 having under 10 employees and Interviewee 1 and Interviewee 6 having under 5.

It is important to note that some of the companies belonged to a concern or oversaw several companies. Interviewee 3 and Interviewee 8 were both coincidentally part of the same concern, whereas Interviewee 4 and Interviewee 5 oversaw several companies.

Majority of the companies interviewed operated within the restaurant/bar industry. Two interviewees also worked with or in a hotel and one interviewee worked in a goldsmith's boutique. Two interviewees, Interviewee 4 and Interviewee 5 had several companies which they managed in some way.

Of the respondents, 5 were male, and 3 were female. While age was not asked from the interviewees, the age range was broad. Interviewee 2 and Interviewee 5 were the only interviewees that could be considered clearly younger than the average, while Interviewee 4 and Interviewee 6 were somewhat older than the estimated average among the respondents. However, at this sample size, few conclusions can be drawn based on the respondents' age and gender alone. Neither the gender nor the age of the respondents was considered during recruitment. The position of the interviewees also varied from co-owners to restaurant managers. However as might be expected, the respondents from the smaller companies were the owners/entrepreneurs or their spouses who also worked in the company. In the larger category the respondents were for example restaurant managers etc. A description of the sample can be seen in table 3.

The interviewees were mainly contacted through the companies' Facebook pages, as this was deemed as the best way of directly contacting those who work with the companies' social media for example. Also, some interviewees were

recruited through personal contacts, however this was a minority. The background research process regarding the choice of contacted companies was conducted through google maps and Facebook for example as this was the most effective way of finding possible companies that fit the description.

Regarding the city size, four size categories are made in order to prevent identification: A meaning population of more than 200000, B meaning 50000-200000, C meaning 20000-50000 and D meaning under 20000. Also, in cases where the company had for example restaurants in other cities as well, the city in which the interviewee worked was used. The city/parish sizes are based on the Finnish statistics center, Tilastokeskus ("Tunnuslukuja väestöstä muuttujina Alue, Tiedot ja Vuosi", 2022).

Additionally, Table 4 shows the google review rating and volume (as of 24.5.2022). Google ratings were chosen as all of the interviewees reported using it for their reviews. As Interviewee 4 and Interviewee 5 oversaw multiple locations, their ratings are listed separately for each location.

For Interviewee 4, the overseen companies are listed as Interviewee 4A, Interviewee 4B, Interviewee 4C, Interviewee 4D, and Interviewee 4E. Interviewee 4A and Interviewee 4B are food delivery companies. Interviewee 4C is a café and Interviewee 4D, and Interviewee 4E are restaurants. As for Interviewee 5, the overseen companies are marked as Interviewee 5A, Interviewee 5B, Interviewee 5C, and Interviewee 5D. Interviewee 5A is a hotel, while the rest are different types of restaurants. All the Interviewee 4 and Interviewee 5 locations are within one city. The number of employees listed in table 3 consists of all these "sub-companies".

The google reviews rating among the interviewed companies were very similar, with only Interviewee 4B having notably lower rating. It should be noted that Interviewee 4B has a low number of reviews (31), making the rating less reliable. The volumes among the companies varied greatly, with the most reviewed having almost 1500 reviews and the least reviewed having only 28.

Table 3 Study sample

Interviewee	Industry	City/parish size	Interviewee position	Gender of the respondent	Approximate size (Number of employees)
Interviewee 1	Restaurant	D	Co-owner/entrepreneur	M	<5
Interviewee 2	Restaurant	A	Local restaurant manager	F	approx. 50

Interviewee 3	Restaurant/bar	B	Local restaurant manager	M	<10*
Interviewee 4	Restaurant/catering	D	Chairman of the Board	M	25-50
Interviewee 5	Travel/restaurant	B	Business director	F	50-100
Interviewee 6	Jeweler	D	Employee (in charge of social media). Spouse of owner.	F	<5
Interviewee 7	Restaurant	B	Co-owner/entrepreneur	M	<10
Interviewee 8	Travel	B	Reception manager	M	10-25

Table 4 Review ratings and rating volumes on Google for the sample as of June 2022

Interviewee	Company rating (out of five)	Number of reviews
Interviewee 1	4.5	221
Interviewee 2	4.5	195
Interviewee 3	4.2	218
Interviewee 4A	4.2	57
Interviewee 4B	3.6	31
Interviewee 4C	4.3	233
Interviewee 4D	4.1	511
Interviewee 4E	4.3	84

Inter-viewee 5A	4.5	911
Inter-viewee 5B	4.6	183
Inter-viewee 5C	4.7	426
Inter-viewee 5D	4.5	103
Inter-viewee 6	4.6	28
Inter-viewee 7	4.3	1457
Inter-viewee 8	4.1	863

4.4 Data analysis

After the interviews were conducted and transcribed, the gathered data was thematically analyzed. As the questions were grouped based on the security action cycle and asked in order from deterrence to remedies, the data analysis was made easier. It is to be noted that as the interviews were conducted in Finnish, the transcripts also were in Finnish, with only the relevant quotes presented in this thesis being translated to English.

Initially, after the interviews were transcribed, the transcripts were analyzed and coded. The coding was done via the commenting feature in Microsoft Word, through which even long sentences were summarized and given the initial code, for example "too small market size". The most important of these summarized sentences were then grouped under their respective themes, which were the four phases of the cycle as well as the second research question. There was a need for additional theming in most phases, such as prevention and remedies as these phases received the most answers. These different themes are roughly visible for example in the subheadings of the preventative and remedial actions chapters in the findings chapter. Answers related to purchasing reviews for themselves was also themed according to whether the interviewee was referring to the negative or positive false reviews, as well as their opinion on them.

The interviewees themselves were also categorized based on some answers they gave, for example had they observed false reviews, responded to them etc. The most significant categorizations can be seen in the different tables presented in the findings chapter.

Theming the opinions on any certain subject was done through a simple five step scale, where "++" was marked to signify strong agreement and "- -" marked to signify strong disagreement. "+" and "-" were also used to signify moderate agreement and disagreement respectively. Indifference or uncertain

answers were marked as "+-". This method was chosen over a numerical Likert scale for example, as it conveyed the attitude towards a matter much more intrinsically than a number at a glance.

5 FINDINGS

Based on the data gathered from the interviews, insight can be gained to **"What actions are companies employing to prevent damages from fake internet reviews and why?"** (RQ1) and what are their viewpoints on the phenomenon. The data also sheds light into the interviewees' opinions about **"What factors deter a company from colluding with fake reviews?"** (RQ2). The different opinions and actions employed by the companies are thematically analyzed based on their category relating to the security action cycle. Answers to these research questions are presented at the end of their respective chapters, figure 4 and figure 5 respectively.

In general, very little to none was consciously done to prevent these damages. These conscious actions were mainly limited to responding to suspected false reviews even though in most cases not responding was seen as the better alternative. This was somewhat expected as there is little that a restaurant or a hotel can do pre-emptively to combat false reviews. However, other factors related to the companies' day to day business were seen to deter false reviews and help them prepare and recover from them.

As none of the companies had experienced false reviews past the individual false review, the remedial actions are mainly based on what would be done in case of a more severe situation. These factors included for example good reputation that would prevent the false reviews from having an effect, not giving any reason for the customer or anyone else to write false reviews, and the rating returning to normal after a negative review campaign. Some actions such as removing comments and arguing with/trying to correct the reviews on the other hand were seen to have a negative effect on top of the attempted corrections.

As for the second research question, the deterrents for buying false reviews were related to ethics and the imagined consequences from getting caught. It is however worth noting that buying positive false reviews was not seen as bad as buying negative false reviews to slander competition. Also, some interviewees did equate buying positive reviews to advertising and thus did not really condemn it very harshly.

When recruiting the interviewees, the companies' sizes were considered to see whether it had an effect. This divided the sample in half: very small, and somewhat larger companies. The size alone did not seem to have much of an effect on the answers. What seemed to have some effect was the field, as the companies that were travel related, took the reviews more seriously and in general were more knowledgeable of them. However, at this sample size it is hardly conclusive. Another factor that seemed to affect the interviewees attitudes towards the false reviews was the position that they had in the company. For example, the owners saw the positive false reviews as more akin to marketing than the others and thus less unethical.

While the interviewees came from several different fields, the common consensus among them is that false reviews are mainly received from individuals who for some reason are not happy with the company. While only Interviewee 8 mentioned that the false reviewers could be paid to write the responses, not much though was given to it during the interview. Also, the only interviewees to mention competition as possible cause for false reviews were Interviewee 4 and Interviewee 5. Interestingly both travel industry respondents (Interviewee 5 and Interviewee 8) mentioned reasons beyond the disgruntled individual, while only one of the non-travel related companies (Interviewee 4) gave a similar reasoning.

The general opinion towards false reviews was negative, however less so for the positive ones. Most interviewees condemned the practice completely mainly due to ethics and the fear of consequences for getting caught. The false reviews were equated to false advertising and fraud for example. Some interviewees on the other hand thought that positive false reviews were just marketing and, in a way, fair game.

While the respondents were actively handling their respective companies' social media and reviews, they generally did not seem to look past their own reviews. Only a few respondents mentioned cases related to false reviews that they had observed. Also, as only Interviewee 3 had observed a positive false review campaign on a local restaurant, the experiences among the interviewees about the subject was limited. This chapter provides an analysis on the obtained results and is loosely structured around the phases of the security action cycle.

5.1 Overview

This chapter provides a general overview about the companies' social media presence as well as experiences and opinions related to false reviews in general. All the interviewed companies reported using Google reviews for reviews. Also, all interviewees took Facebook and other social media channels into consideration when talking about reviews and reputation management as they also received reviews there. The companies providing services in the travel industry (Interviewee 5 and Interviewee 8) reported using online travel agents (OTA) such as Booking.com, expedia, trivago and TripAdvisor. Other social media and review channels were also in use by some companies such as the food delivery app

Wolt. Table 5 lists the platforms related to reviews and reputation management used by the companies as reported by the interviewees. Other ways of receiving feedback mentioned were interviewing the customers before leaving, private feedback channels and online chat forums where the interviewee had seen discussions about their company.

Table 5 Review platforms used by the companies

Interviewee	Platforms in use by the company
Interviewee 1	Facebook, Google reviews
Interviewee 2	Facebook, Google reviews, Wolt
Interviewee 3	Facebook, Google reviews
Interviewee 4	Facebook, Google reviews, TikTok, Twitter
Interviewee 5	Facebook, Google reviews, OTA
Interviewee 6	Facebook, Google reviews
Interviewee 7	Facebook, Google reviews
Interviewee 8	Facebook, Google reviews, OTA

The aforementioned platforms were in regular use regarding the reviews, with answers ranging from daily use to weekly use, with Interviewee 4, Interviewee 5, and Interviewee 8 having regular meetings about reviews, whereas the others acted more on ad hoc basis.

Two of the companies, Interviewee 4 and Interviewee 7 reported having a professional social media manager. In the case of Interviewee 7, the social media and reviews were in completely outsourced to a professional company while Interviewee 4 had a social media manager within the company. Interviewee 7 nor any of the co-owners of the restaurant were not involved in any social media personally and only communicated with the social media manager when their input was required. Interviewee 7 had no idea what their reputation was online and reckoned that it was not good. However, according to google reviews the company had a 4.2-star rating. When asked about the managing of social media, Interviewee 7 answered *"So, we have an outside person in Helsinki who handles it. We do not even have the credentials to our platform so that we cannot comment anything, so that we don't get provoked."* (Interviewee 7). Interviewee 4 on the other hand worked together with their social media manager and regularly held meetings to discuss reviews for example.

In general, the influence of reviews and false reviews were acknowledged by the interviewees. For example, Interviewee 1, Interviewee 4, and Interviewee 7 saw online reviews akin to online marketing and Interviewee 8 mentioned the reviews being able to provide information about a company that it does not disclose normally. *"(…) the features are often better presented in the reviews than on the hotel's own pages, obviously because no-one advertises that they don't have this or that."* (Interviewee 8). Interviewee 7 also mentioned that social media and reviews are such a strong force, that when it works in their favor it's amazing but when it

works against them it is very bad. For this reason, Interviewee 7 chose to not to be involved in the process at all personally.

… it is an amazing marketing channel when it works but when it works against you it is many times easier to just not be involved. Either you're in it or straight up don't know anything about what is happening there. (Interviewee 7)

When asked about a sudden surge of negative reviews, akin to a false review campaign, almost all respondents felt that the repercussions would be severe. Only Interviewee 1 was very confident on being resistant in case of numerous false reviews.

Even if I get a hundred negative reviews, it means nothing. And that's not even an arrogant claim, because in the end it boils down to how stable the situation is, meaning how the actual customer sees the place in terms of quality. (Interviewee 1)

Most of the interviewees had had some sort of experience with false reviews related to their company, however most stated that their experiences were limited to only a few reviews or comments that they had observed. Also, some false reviews had been personally observed or suspected in other companies.

Interviewee 1 had the most experience with false reviews among the participants and was the only one with some sort of plan in place for dealing with false reviews. When asked about experiences with false reviews he replied: *"They do always exist and mainly they are from the people that are somehow jealous."* (Interviewee 1). Interviewee 1 also expressed his opinion on false reviews and concerns over how commonplace they seem to be by saying *"It's all in all very childish and there is very much of that that you can buy likes and followers."* (Interviewee 1).

Some of the respondents felt that recent events had affected the false reviews. For example, Interviewee 2 felt that the war in Ukraine had sparked hate comments and reviews towards Russian owned, or themed restaurants: "Perhaps these false reviews have surfaced because of the war in Ukraine. The restaurants are given reviews where one's opinion on the war is stated." (Interviewee 2).

Interviewee 3 mentioned that he had observed false reviews on other restaurants/bars due to their stance on COVID-19 protocols. *"Especially in a place like [a finnish celebrity chef]'s restaurants, people left many 1-star reviews just because they didn't agree with his opinions about COVID related things."* (Interviewee 3). Interviewee 3 had also noticed hundreds of 5-star reviews for a nearby restaurant that had not been opened yet and suspected that they were from a religious group that the restaurant was affiliated with somehow.

… I once noticed that there were lots of false reviews when [the restaurant in question] had an article written about on [local newspaper] and there was a link, so I followed it to their Facebook page. And the place was not even open yet, it said that opening in a month or so and there were hundreds of 5-star reviews at the time. I did a bit of background research at that point and found out that it's owned by

some religious group and apparently had their own people vote on the unopened restaurant very positively. (Interviewee 3)

Interviewee 3 also mentioned that their company and he personally had received hate comments of Facebook because of the company's decision to use the vaccination certificate. Some false reviews were also received from drunken customers who were thrown out of the bar/restaurant.

Interviewee 8 had observed false reviews on foreign webstores. He stated that *"For example on these foreign webstores and so on, there can be the exact same text on the reviews or some other shoddy stories. Those have come up every now and then."* (Interviewee 8).

Other reasons stated for the existence of false reviews mentioned were exaggerated negative experiences, which were mentioned by Interviewee 2 and Interviewee 8. Interviewee 8 also mentioned that some reviews might be considered false due to the perspective difference between the reader and writer. *"One man's crappy is other man's luxurious and the other way around so it's not meant to be fake even though there's a discrepancy, which is something to you should always consider"*. (Interviewee 8).

Table 6 shows the motives that the interviewees felt were the reason for deliberate false reviews in general.

Table 6 Interviewee's opinion on the general motives behind deliberate false reviews

Interviewee	Motives for false reviews
Interviewee 1	Jealousy, spite
Interviewee 2	Frustration from bad experience, targets nationality
Interviewee 3	Quick boost in visibility, difference in opinion about something unrelated
Interviewee 4	Competition, personal grief/grudge
Interviewee 5	Mischief, slander competition
Interviewee 6	Boost own business or slander competition
Interviewee 7	Slander in general, clout
Interviewee 8	Too much time on the writer's hands, paid to review

When asked generally about what their opinion about false online reviews were, all of the eight respondents responded in a negative manner. Some of them condemned the practice harder than others. For example, Interviewee 4 thought that the phenomenon of false reviews was commonplace nowadays and nothing special by saying *"Yea it seems like it's part of everyday life nowadays, luckily, it's marginal in our case but it has happened."* (Interviewee 4).

False reviews were seen as very detrimental for new companies by some interviewees. Interviewee 1 for example mentioned that they were much more active in social media in the beginning and the poor reviews hurt much more

than they do now, regardless of whether they are true or not. Interviewee 7 also mentioned that a new company getting constant false reviews would likely seriously affect their success. *"If you have been open for a month and all you get is negatives [reviews], I don't think the cash register is going to like it."* (Interviewee 7).

Interviewee 2, overseeing the newest location of the restaurant, also mentioned that they would feel the effects of false reviews more than their other restaurants because the locals do not yet know the place well. When asked about how their reputation would hold in case of a "busload of negative reviews", they responded *"I would say it would hold. Or let's say that here in [City], because we've only been here for a year and a half (…) here it could have a stronger impact."* (Interviewee 2).

All in all, the respondent's first-hand experiences about false reviews were all related to individual disgruntled customers or to people who had something personal with the company or interviewee. Some respondents had observed false reviews on unopened restaurants Facebook-pages, on famous restaurants pages and on foreign webstores. The consensus among the respondents seemed to be that Finland is too small for any larger false review campaign to work. Despite acknowledging the effect of false reviews, none of the interviewed companies had publicly discussed false reviews. Also, only a few had internally discussed false reviews specifically, while reviews in general were an everyday subject.

5.2   Deterrent actions

Deterrence in terms of actions done by the companies did not really exist among the interviewed companies as for example warnings or other preventative publications etc., were not thought of and were seen as a bad idea. This was expected as it would make no sense from the companies' point of view to pre-emptively to start deterring false reviewers especially as many felt that by getting provoked, the problem would only get worse. Also, as Interviewee 7 stated, retaliating is often a sign that something is wrong.

However, there were two factors that cannot really be attributed to the other three categories: firstly, pre-built reputation, which in essence prevented the false reviews from having an effect and thus acting as a deterrent and secondly, small market areas. Latter of which increased the chance for getting caught from buying reviews and thus relating to the detective and remedial phases of the cycle.

The first factor that could be considered to be in this category is pre-built reputation. Interviewee 1, Interviewee 2 and Interviewee 7 felt that their pre-built reputation worked as a deterrent for the possible writer of false reviews. For example, Interviewee 7's restaurant had been in the same place for 25 years, while Interviewee 1 thought that the only opinion that really matters is the actual customer's, which spreads through word-of-mouth *"If someone's friend likes something, then when their friend also sees it, which makes it clear marketing."* (Interviewee 1). On Interviewee 1s part this is likely stemming from the company's aggressive marketing that he mentioned. As for Interviewee 7, the company had been there

for 25 years already and had a good reputation and plenty of reviews as presented earlier in table 4. As for Interviewee 2, the explanation may be related to the prior reputation from other locations.

As regular customers of the companies have first-hand experiences and knowledge of the companies, it can be expected that they tend to read less reviews about them in the first place. This makes the outsiders the main audience for online reviews. This in a way nullifies the benefits from the regulars being able to see past the slanderous reviews. However as most of the interviewed companies were in the restaurant industry, the main customer segment is likely the locals, who already know the place. This was for example brought up by Interviewee 5.

> *During wintertime, we have lots of company groups and other regulars. The regular customers of course do not write reviews there [on TripAdvisor] and they don't search it for restaurants either. The locals know where to go even without TripAdvisor. (Interviewee 5)*

The second factor in this category is the market size. Many of the respondents stated that their respective market areas and even Finland in general are too small for false reviews to be effective as word spreads around quickly. Interviewee 5, Interviewee 6, Interviewee 7 and Interviewee 8 mentioned that their respective market areas were too small for any bigger false review campaigns to work. Interviewee 5 and Interviewee 8, both being hotels, stated that Finland is too small as the hotel industry is too small for such campaigns because word gets around quickly when people know each other. *"It's difficult to imagine that in a market the size of [city of Interviewee 5]. I mean in general this field is quite small in Finland, meaning you would get caught pretty quickly if you tried something like this"* (Interviewee 5). When asked to specify why this was the case Interviewee 5 responded with *"Well, the circles are small. Would it stay a secret in your workplace that you have written reviews for the company (…)?"* (Interviewee 5). Interviewee 7 had a similar response, stating *"Yes that's how it is. This place [meaning city] is too small. Basically, everyone in the restaurant industry knows each other here. It's like a small family (…)."* when asked whether it would be obvious who ordered a false review campaign. He also thought that in bigger cities, the industry is much more *"faceless"*, which could make this type of campaign viable. Interviewee 6 had the same idea as she thought that in small cities, the businesses often become personified

The fact that word gets around quickly and that the different companies are well connected, and close knit was seen as a reason for why false reviews don't really exist in Finland by Interviewee 2, Interviewee 5, and Interviewee 7. This type of non-cutthroat environment in terms of competition is also a likely reason as to why there seems to be no stress about negative false reviews beyond the occasional annoyed customer. While the absence of any serious cases of false reviews cannot completely be attributed to these factors, they surely are a part of it, and a likely a major reason for the lax preparations.

5.3 Preventative actions

Few actions were consciously done to prevent false reviews from having an effect beforehand among the interviewed companies. This was quite expected as there is little that an individual company can do to prevent false reviews. The preventative actions were attributed to the review platforms, who were imagined preventing false reviews through high posting costs and by having filters detect the fakes before they are published. From the companies' side, not giving the customer any reason to get angry was the main factor that was thought to prevent false reviews. Also, some so-called pre-screening actions were seen to prevent false reviews. Figure 3 attempts to visualize the reasons for the lack of preparation against false reviews based on the interviewees' answers.

5.3.1 High posting costs

Some interviewees thought that the high posting costs present on the different platforms would act as a preventative and for example Interviewee 5 hoped that the practice would be more widespread, however a verification of purchase in order to review a restaurant would likely be difficult implement. Interviewee 5 also felt that Google maps and TripAdvisor were more prone to false reviews due to their low posting costs than websites such as Booking.com and Expedia, which required proof of purchase before being able to review.

Related to the high posting costs, some respondents thought that requiring a real ID to post reviews would deter potential false reviewers. Interviewee 3 felt that by having to write with your own name, people would be less inclined to write politically charged false reviews "(…) I wouldn't want my name out there for everyone to see that I was wrong about something. Well not necessarily wrong but against the common consensus." (Interviewee 3). Interviewee 3 suggested that a strong identification should be required to improve the quality of discussion on some platforms but on the other hand saw it as a slippery slope of sorts. Interviewee 4, Interviewee 6, and Interviewee 7 also felt that using one's real name would serve as a deterrent and should be required when posting reviews. This type of strong identification as a preventative would obviously fall on the platforms to enforce. While posting with a real name and picture makes one's reviews more reliable, it does not come without its privacy concerns for example. The respondents suggesting this were on the older side in terms of age, which could imply that they are less aware of the possible risks involved.

5.3.2 Pre-screening

Some thoughts about stopping the false review before it is posted were also brought up in the interviews. Interviewee 2 and Interviewee 7 for example trusted the platforms to have some sort of filters that would catch some the comments or reviews before they were posted on the website. Interviewee 2 specified

that these types of filters might deter the false reviewer depending on their motivations for posting:

> *(…) and then there would be some banner that said that "we detected something something in your text" I believe it would make someone think twice… but it would only probably affect the indignant customer. If someone wants to deliberately false review, no amount of moderation or other can prevent it. (Interviewee 2)*

However, Interviewee 2 and Interviewee 6 though that moderation of the comments i.e., removing them after they have been posted would not serve as a deterrent for most.

One of the few preventative actions mentioned by the interviewees was brought up by Interviewee 8, who mentioned that by interviewing the hotel visitor when checking out, he believed that less bad reviews would surface. *"We interview about 90% of our customers when checking out. By getting the good feedback and the bad feedback directly (…) only a fraction of it will surface out there for everyone to see."* (Interviewee 8). In this case the would-be false reviewer could vent their frustration during the interview and not on the internet. However, the purpose of the interview in their case was to gather feedback and not deter possible false reviews. This type of practice could however be difficult to implement for restaurants and even more so for smaller kiosks and bars for example due to sole initial service interaction before the consumption of goods.

5.3.3   Keeping the customer satisfied

Not doing anything provocative was seen as the best way to prevent the individual from being upset in the first place. Some interviewees replied and others implied that by just doing good work and leaving the customer satisfied, the individuals would have less reason to post negative false reviews. As all the interviewed companies had good ratings on google reviews, this is likely working. This is also likely closely connected to the pre-built reputation presented in the deterrent actions chapter.

Interviewee 2 for example mentioned that by doing a good job in the first place and not giving any reason for the customer to get upset, false and negative reviews in general would be avoided. *"When you don't give the would-be false reviewer any reason to get provoked or anything (…). We have a good reputation, and we do a good job and so on, which gives no reason for someone to get upset."* (Interviewee 2).

As the consensus among the respondents was that false reviews are mostly motivated by personal grudge or bad experiences, which then are expressed exaggeratedly or with some redactions to make their experiences seem worse, keeping the customer satisfied allows them to thwart this risk.
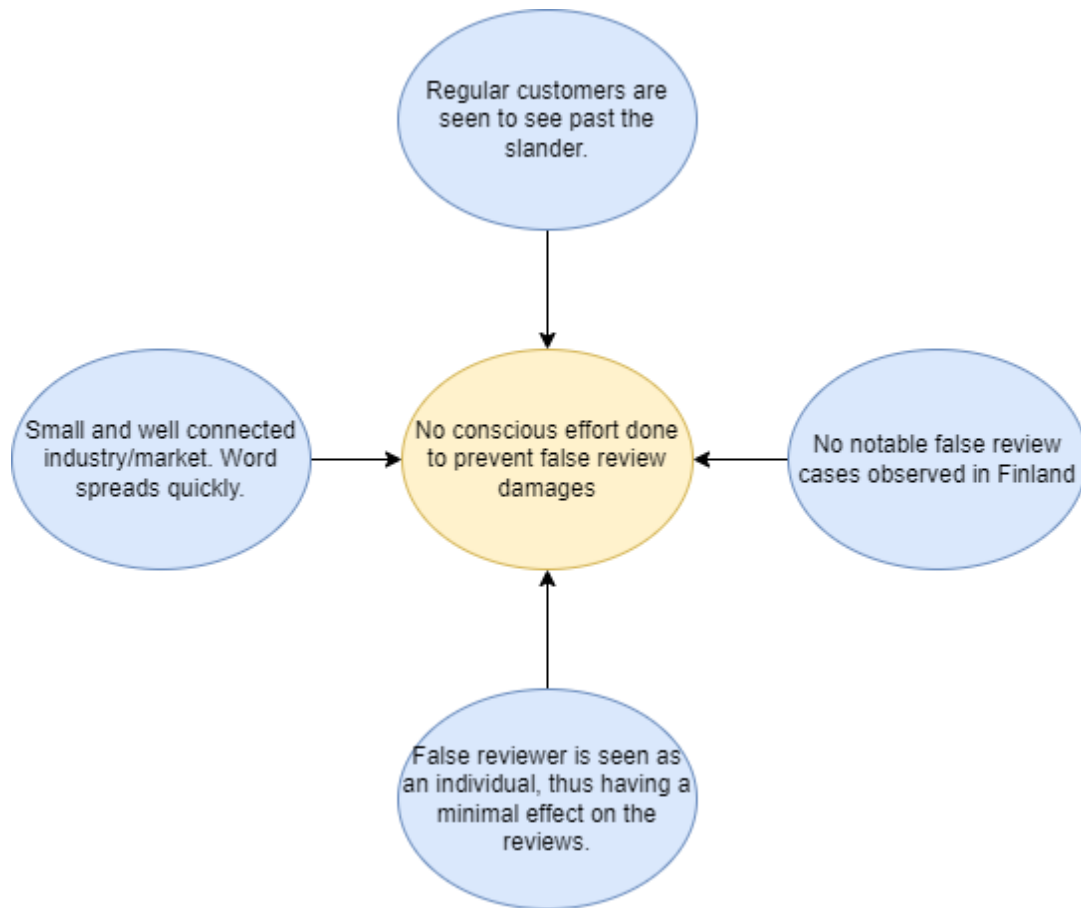
Figure 3 Factors contributing to the lack of preparation against false reviews from the interviewee's viewpoint

Some felt that there is not really a way to stop false reviews from happening. For example, Interviewee 7 when asked about whose responsibility is it to stop the false reviews responded: *"I'm not sure if anyone can prevent them. (…). I think the responsibility lies within the one posting the content online but currently it does not seem to exist (…)."* (Interviewee 7). Interviewee 6 also was doubtful whether false reviews can be stopped when asked about what the platforms are doing to prevent false reviews. These opinions could be attributed to lack of knowledge about false reviews as most of the older respondents mentioned not being as tech savvy.

While many of the interviewees had experienced false reviews in some way, they were mainly from individuals. Their motivations were imagined to be based on poor experiences and/or personal grudge with the owner and in the end having minimal effect on the company's rating or reputation. This is likely a major factor on why the subject of false reviews has received such little thought in the companies in terms of prevention.

5.4 Detective actions

The detective actions come in to play once the false review has been written. Most respondents felt that in order to distinguish false reviews from the real ones, you must have some previous information about the company or other relevant experience. Factors that were mentioned to cause suspicion the most were anonymous reviewer and factual mistakes. In general, this phase included the least deterrents and actions as the consequences from getting caught are attributed into the remedies phase. The only action in this category is the detection of the false reviews. However, meetings related to reviews in general can also be considered to belong in this category. As mentioned earlier Interviewee 4, Interviewee 5, and Interviewee 8 mentioned having meetings about reviews on the regular, while the others had meetings when needed. In essence, this chapter presents the interviewees opinions on how well they think the reader is able to detect the false reviews as well as what factors cause themselves to suspect foul play.

When asked how well they think the reader can detect false reviews online, the interviewees had somewhat similar responses. The interviewees also elaborated on what factors related to the reviews would cause suspicion. Most respondents thought that detecting false reviews required previous information about the company in some way. Only Interviewee 5 thought that the readers would easily spot the false reviews in general. Interviewee 1, Interviewee 4, Interviewee 6, and Interviewee 7 thought that only the regular customers would detect the false reviews as they know the quality of the company from before. Interviewee 1 also added that *"people can distinguish a false review based on how it's written. Often real criticism is written properly while false reviews are slanderous messages for the sake of slander"*. (Interviewee 1). Interviewee 5 also mentioned that descriptive reviews are more often truthful.

In general, the notion was that people would be able to see past the false reviews, mediated by different factors such as age and experience with reviews. This confidence on the ability to detect false reviews is likely misplaced as covered in the literature review chapter about false review characteristics. A possible factor boosting this confidence could be the information asymmetry existing between the interviewee and the reader about the company. As the employee might detect false reviews based on factual mistakes and context, they might also unintentionally expect the reader to do the same.

Detection based on the reviewers themselves was also a common answer. Whether it was confirming the reviewer through reservations, by determining the truthfulness of the profile based on the displayed name and picture or just by knowing or identifying the reviewer personally. For example, Interviewee 2 mentioned that several anonymous profiles posting reviews that are against the consensus are likely fake and, in such case, people would be able to detect them.

Interviewee 8 on the other hand thought that with enough experience with online reviews, one might be able to detect them. Younger people were thought to perform better in detecting false reviews due to their media skills according to Interviewee 3. Related to this, it is possible that Interviewee 5, being one of the

youngest respondents, might have overestimated the media skills of the older population when stating that most can detect the false reviews. Interviewee 3 and Interviewee 5 responses are seemingly similar and give the idea that general experience on reviews is also useful when determining truthfulness and not just knowledge of the specific company.

The effectiveness of the false reviews was attributed partly to the single numerical rating not providing enough information to the reader. When asked about people's performance on detecting false reviews Interviewee 3 felt that false reviews are in part so effective because of the platforms often initially showing only the average rating.

> *I would suspect that people perform quite poorly. In that sense it can be a working strategy to buy false reviews as at least I don't inspect them too closely. When going to a new place I might take a look at the overall rating at best. (Interviewee 3)*

An essential factor related to detecting false reviews, bimodal reviews, was not mentioned by any interviewee. In general, the overall look of the rating spread was not talked either other than what Interviewee 3 said. It is likely that as all of the companies took Facebook comments into consideration, the weight put on the numerical rating itself was less as the focus was more on the written content of the comment/review.

However, a sudden burst of negative reviews was mentioned by three interviewees but not in relation to the spread of the rating. Table 7 gives a general overview on the interviewees' views on the readers' abilities to spot false reviews.

Table 7 Interviewees' opinion on readers' ability to spot false reviews

Company	Opinion on readers ability to spot false review
Interviewee 1	Only regulars with experience about the company. Also, some clear cases where it is obvious.
Interviewee 2	In clear cases only.
Interviewee 3	Poor in general, younger people perform better.
Interviewee 4	Only regulars with experience about the company.
Interviewee 5	Most can distinguish.
Interviewee 6	Regulars and people with information about the company can.
Interviewee 7	Somewhat good, regulars even better.

Interviewee 8	People with experience on reviews can.

When asked about factors that would cause them to suspect false reviews, a broad range of answers was gotten. The most common factor relating to the reviews themselves was the lack of real name or profile picture on the reviewer. This was mentioned by all respondents when asked about the subject. "*(…) if I saw some strange reviews and their name was not real or they had no picture or some strange picture. It would make me think, the anonymous profile specifically.*" (Interviewee 2).

Another common way for the respondents to determine the truthfulness were the factual mistakes in context. For example, Interviewee 5 and Interviewee 8 could quickly see via reservations whether the reviewer had actually visited their restaurant/hotel "*People do table reservations with their name, so of course I will recognize them if I have looked at the reservations.*" (Interviewee 5). On the other hand, Interviewee 1 had noticed negative comments about his establishment on a popular online forum and recognized a few of the writers "*At some point I looked at Suomi24.fi and I recognized a few of the commenters, and for example my neighbor was one of the shit talkers, which says a lot about him.*" (Interviewee 1). This type of ability to confirm the truthfulness based on external factors was mentioned or implied by Interviewee 1, Interviewee 2, Interviewee 3, Interviewee 5, Interviewee 7, and Interviewee 8. While this type of false review detection works for the companies as they typically know what has happened, the reader of the reviews likely does not. This would logically feed the information asymmetry between the two parties, which was already discussed earlier in this chapter. Other minor factors mentioned by the interviewees included a sudden burst of reviews, where many positive or negative reviews would emerge in a short timeframe and a slanderous writing style. Also, other factors related to the review were the valence i.e., the rating, language used, and similarity to other reviews (repeating phrases). Table 8 gives an overview of these factors.

Table 8 Factors that cause false review suspicion in reviews

Interviewee	Factual mistakes	Suspicious reviewer	Sudden burst	Writing style	Review valence	Similarity to other reviews
Interviewee 1	X	X	X	X		
Interviewee 2	X	X			X	
Interviewee 3	X	X	X	X		
Interviewee 4		X	X			
Interviewee 5	X	X				

Inter-viewee 6		X				
Inter-viewee 7	X	X				
Inter-viewee 8	X	X				X

Getting caught and the subsequent consequences were seen as a deterrent factor by Interviewee 3, Interviewee 5, Interviewee 6, and Interviewee 8. They expected the risk of getting caught to deter the false reviewers and the buyers. However, some exceptions based on motivation was made by Interviewee 8, who thought that a motivated writer would treat it as a challenge and that they would have to do better if they got caught. As for the buyer, Interviewee 8 thought that getting caught would serve as a deterrent.

5.5 Remedial actions

Most of the companies' own actions in case of false reviews would fall under the remedies category, as there is not much they can do pre-emptively. However, the current response strategies that the companies had in place for false reviews were slim. When asked about how they would react to a more severe false review surge, the interviewees came up with several different actions that they imagined would help them weather the storm.

Replying to comments was seen as a double-edged sword, as it could cause more harm depending on the situation. However, some saw it as a good remedy in situations where the review was clearly fake for example. In general, as the individual review does not have much impact on its own, many chose to just leave them be. The second remedy was more of a passive one. Closely related to the pre-build reputation, the rating returning to normal after a while was seen to deter some false reviewers depending on their motivations.

Using review platforms as a part of remedy received mixed opinions. On the other hand, the respondents saw no harm in trying to contact them in order to remove reviews but on the other hand many were doubtful whether this would amount to anything. Other criticism was also directed towards the plat-forms presented in this chapter as well as in general in the interviews.

Removing suspected false reviews received opinions similar to commenting on them. While removing slanderous content online was seen as a good idea, some elaborated that too much censorship could backfire on them. Lastly this chapter looks at punishments and other consequences that a company buying false reviews would face. The remedial actions that would be employed are presented in table 11 at the end of this chapter.

5.5.1 Replying to false reviews

While the other respondents came up with ideas when asked about how they would react in case of their company getting repeatedly false reviewed via a false review campaign, Interviewee 1 was the only company that reported some sort of plan in place in terms of false reviews. Interviewee 1 had a strategy to deal with false reviews that was almost akin to crowdturfing, where he would urge his friends to challenge the review or comment.

> *I have a couple of friends who you could call basically professional pricks. If some rando starts talking shit on the grill's page especially if it makes no sense (…) so I sent in these professional pricks who can through clever writing, verbally murder them and a few comments later the woman who started the conversation deleted her own comment. (Interviewee 1)*

Somewhat related to this, Interviewee 4 had also had outsiders and customers defend their company from false reviews, however in this case it was not asked for. "*(…) the false reviews have of course been public and often times other people have then commented on them.*" (Interviewee 4). When asked to elaborate, Interviewee 4 said:

> *People have defended us. Someone had complained that their portion had a maggot in it and attached a picture (…) and someone who knew the reviewer suggested that maybe the maggot had gotten into the food during the delivery instead of the cooking process (…). (Interviewee 4)*

For Interviewee 1, response as a remedy provides to the deterrence feedback loop by having public responses to the fake reviews displaying that the content of the message will be challenged by someone. The experience with false reviews, combined with experience with marketing is a likely contributor to the almost overflowing confidence that Interviewee 1 had regarding being able to deal with false reviews. As for Interviewee 4, their reputation in the eyes of the other customers may prompt an answer from someone reading the reviews.

Responding to suspected false reviews was seen to lessen their effect on the subsequent reader according to Interviewee 8. Interviewee 8 had had a few relevant experiences during the last two years. He felt that responding to a written review is important as the writer has spent time writing the comment in the first place and that the company's response help the subsequent readers of the comment and its response.

> *(…) if there's a review, be it any kind, I will write the response in a way that it reaches the others who read the comment. For example, I may include some self-promotion or other promise of better service. (Interviewee 8)*

Interviewee 4 on the other hand felt that it was their word against someone else's, implying that it would be difficult for the reader to determine which one was telling the truth. Also, Interviewee 4 generally wasn't sure how they could respond in such situation. This points to at least Interviewee 4 acknowledging that their reply would not have much weight in the comment sections.

Responding to false reviews as a remedy in general received a mixed bag of answers. While many stated that it would depend highly on the circumstances, most of who would respond, had responded and those who wouldn't respond, had not. The idea behind this is likely the fact that the writer of false reviews is imagined as an individual and not as a professional astroturfer or crowdturfer. By responding, the company would further provoke the individual, which in turn could escalate the situation. Whereas by not responding the false review would still be there, but on its own would have little effect on the readers' opinions.

The most common, current plan of action regarding false reviews was to just let them be. The interviewees felt that there are so few false reviews that it is better to not do anything. Some felt that responding by correcting factual mistakes was a good idea in some clear cases while others felt that responding to a negative review might make them look bad. For example, Interviewee 7 described such situation:

> *Let's put it like this, if I would go and try to correct slanderous reviews, the customers might get a feeling that is there something more to this, why such an aggressive response style. Often a retaliation is like a sign that something bad has happened. (Interviewee 7)*

Interviewee 1 felt that by replying to the provocative false reviews, the writer's actions were validated and thus sometimes it was better to leave them alone. Interviewee 1 did not also particularly care about any negative, non-constructive criticism regardless of the source. The sentiment of not responding was shared by Interviewee 2, Interviewee 3, Interviewee 4 and Interviewee 7 as well, who felt that by getting provoked, the problem would grow even larger. As individual false reviews are not very influential, most respondents likely chose to ignore them.

Interviewee 3 and Interviewee 5 mentioned that responding properly and systematically to the online reviews in general would require more manpower. In the case of Interviewee 5 it was a company policy to not respond due to this. If given more resources, Interviewee 3 implied that they would react to a potential surge of false reviews by responding but he saw it a little bit risky.

> *I wouldn't go about hiding it. I would try to respond to them and discuss about the matter. However often it escalates very easily making it very difficult. (…). I remember a few cases where someone got shat on from every direction and I can't remember any of them emerging as a winner from it. (Interviewee 3)*

Interviewee 1 specifically mentioned not responding being a deterrent for false reviewers when asked whether not responding discourages the false reviewers. *"It absolutely discourages them. If you think it this way that they are bullies and you don't care about their actions, at some point they will get bored. However, if you get provoked, they continue."* (Interviewee 1).

Table 9 shows whether the interviewees would respond to reviews that they felt were fake. While most would not respond normally, some exceptions were mentioned regarding situations when the truthfulness of the review is unclear. When prompted about how they would respond if they had to, all respondents focused on correcting factual mistakes.

Table 9 Would the interviewee respond publicly to a review they deemed false

Interviewee	Has responded	Would respond in some cases	Would not respond
Interviewee 1	X		
Interviewee 2			X
Interviewee 3	X		
Interviewee 4			X
Interviewee 5			X
Interviewee 6		X	
Interviewee 7			X
Interviewee 8	X		

In case of a false review surge, most respondents would seek help from their colleagues or experts when determining how to proceed. Interviewee 5 and Interviewee 8 mentioned that they would make a public statement that they had been the target of a false review campaign.

- Interviewee 2: would have a staff meeting about how to proceed.
- Interviewee 3: would ask help from the concern.
- Interviewee 4: would have a meeting with social media manager.
- Interviewee 5: mentioned that it would be handled around her level in term of hierarchy and as a group effort.
- Interviewee 7: would leave it to their social media professional.
- Interviewee 8: has a three-person group that deal with reviews.
- Interviewee 1 and Interviewee 6 both being the sole person in charge of the social media and reviews, would act on their own.

Interviewee 5 and Interviewee 8 seemed to be more knowledgeable about reviews than most other respondents. They, for example had more review platforms in use, they emphasized the use of reviews as feedback more than others and had contacted platforms in an attempt to remove comments they deemed false. They were also the only ones together with Interviewee 4 who regularly held meetings about the reviews. Interviewee 5 and Interviewee 8 making a

statement in case of a false review campaign aimed at them could indicate that they are perhaps more concerned about their reputation than the other respondents. They were also the only ones to suggest fines as a punishment which is also a likely indicator that they would take the situation more seriously than the others.

This potentially points to the hotels taking reviews more seriously than the restaurants. The reviews might be more influential for the hotels as mentioned previously by Interviewee 5 who stated that locals know where to go without TripAdvisor reviews. On the other hand, both Interviewee 4 and Interviewee 5 oversaw several companies, which could be another reason for the more organized meetings about the reviews. This sample size however cannot provide any clear results on this finding with certainty but could act as a possible impetus for future research.

5.5.2   Rating bouncing back

Having a good rating and continuing to receive good ratings was seen as a deterrent to some extent. When asked about a sudden surge of false reviews, most interviewees believed that their rating would drop temporarily but by continuing business as usual, it would return to normal after a while. Only Interviewee 3, Interviewee 6, and Interviewee 8 were not sure whether this would happen.

The rating bouncing back was seen as a deterrent factor in some cases depending on the reviewer's motives. For example, Interviewee 2 thought that disgruntled individuals might realize the futility of their possible rant and by extension be deterred from posting a false review.

> *If it was some competitor, perhaps not, as if the intention was to*
> *cause harm, for some, the momentary drop in rating might be enough.*
> *But if it's someone who's just angry for some reason it might make*
> *them think twice. (Interviewee 2)*

The results about the rating bouncing back were positive. All interviewees though that the rating returning to normal after a while would serve as a deterrent at least to some extent. Interviewee 2 thought that a short drop in competitors rating could be enough for some however, thus might not deter them. Interviewee 6 and Interviewee 8 also thought that the effectiveness of this was dependant on the writers'/buyers' motivations and personal willpower to continue the false reviewing. "(…) *if they noticed that the rating would bounce back, would they be bothered to try again and again. I don't see it really plausible for someone to continue indefinitely.*" (Interviewee 6).

In general, the rating returning to normal after a while acting as a deterrent was subject to the reviewers' motivations. As a more professional entity with the intention to lower the rating could be satisfied with even a momentary drop in the rating whereas the individual would be deterred when faced with the futility of their actions. It is however questionable to assume that an individual has the intent to cause a drop in the rating with their false review and instead their actions are likely more focused on the written part of the review. In general, the

bouncing back as a deterrence was seen to be dependent on the writer's tenacity to continue their efforts, which in case of a campaign could be costly for the buyer.

5.5.3 Platforms as a part of remedy

When asked about contacting the reviews platforms about possible false reviews, the responses were quite unanimous. Most felt that reporting comments on the platforms would not amount to anything, however they would at least try. Most would only contact the review platform in cases where the reviews were clearly false, for example in case of a campaign targeted towards them. For example, when asked about reporting comments, Interviewee 4 commented *"Probably in cases where the situation was clearly a witch hunt (…)."* (Interviewee 4). As for the apathetic expectations regarding the platforms, Interviewee 5 thought that *"(…) To get a hold of them can take days. And like I said, they likely don't give a crap about some small company here in [Interviewee 5 city]."* (Interviewee 5).

Interviewee 5 and Interviewee 8 though that the review platforms should be in a closer cooperation with the companies regarding false reviews. They both thought that the companies should have a stronger say whether the review is fake but Interviewee 5 for example acknowledged that this could be easily abused by the more dishonest companies. Interviewee 5 also thought that more platforms should have high posting costs. Interviewee 8 on the other hand hoped for better review system in general.

> *(…) some of the platforms have a system that you can give 10, 7.5, 5 or 2.5 and the category is a bit so and so. It should be divided into smaller ones. The rating scale should be more flexible (…). I understand it's easier for the customer to fill out a short review, but they would be more useful for us if there were more free text fields and not just the customer ticking some boxes and giving a numerical rating.*
> *(Interviewee 8)*

It is likely that only Interviewee 5 and Interviewee 8 elaborated more on the review platforms as the others did not employ them. It is likely that these platforms are less popular among restaurant users because it is likely easier for the customer to search for nearby restaurants through google maps for example and get the reviews from there.

5.5.4 Deleting and removing comments

Some of the interviewees talked about removing the comments and false reviews from the public eye. In general, the subject was seen as a double-edged sword. While Interviewee 5 stated that she had removed inappropriate comments from their Facebook page and in general hoped that it would be easier to remove them from platforms, she also acknowledged the problems arising from constantly removing comments. *"If I would remove the same persons comments all the time, they might start asking questions and make it a bigger deal."* (Interviewee 5).

Interviewee 3 on the other hand saw removing the comments difficult, as it is very hard sometimes to distinguish what is fake and what is not. He also thought that being able to remove comments would undermine the purpose of the reviews: "*I don't think companies should be able to remove the comments. It is wrong because it makes the reviews just stupid if you can take the bad one out and leave the good ones in*" (Interviewee 3).

Interviewee 7 also thought that companies should be able to remove obviously fake content more easily. He suggested a third party as a "judge" for determining the case but realized that the number of reviews would make it impossible. Interviewee 7 reflected on this through a following situation that had happened in their restaurant.

> *The customers sometimes do some incredibly stupid stuff. They are typically underage, with two guys doing something while the third one is filming. If you do something or yell at them, the video will be edited to make you look bad. Same thing in schools (…). Then it will be posted on social media and then it finds its way into the evening tabloids (…). (Interviewee 7)*

A different way of removing comments was employed by Interviewee 1, who had his friends challenge the messages in an effort to get the commenter to remove their text. Interviewee 1 did not mention any negative aspects to this type of action.

All in all, removing comments/moderation serving as a deterrent did not receive much support. When asked about whether this deters the writer, Interviewee 6 thought that "*someone who wants to personally slander will find another way.*" *(Interviewee 6).* Interviewee 2 and Interviewee 3 also though that moderation of the comments i.e., removing them after they have been posted would not serve as a deterrent for most. Interviewee 3 thought that moderation of the reviews or comments would only deter the paid reviewer. Interviewee 7 also implied this by stating that "*(…) the removed comments tend to come back quite fast*". *(Interviewee 7).*

This once again likely points to the idea that the false reviewer is seen as an individual and thus if they want to slander, they will find a way. On the other hand, more serious cases could be dealt through this, as most respondents would at least try to contact the platforms in case of an obvious false review campaign.

5.5.5 Punishments and other consequences

All the interviewees who were asked about whether false reviewing should be punishable (Interviewee 3-Interviewee 8), answered positively, however the severities varied for example due to the difficulty of determining the purposefulness. When asked about possible punishment methods to false reviewers the interviewees mainly suggested bans on the platforms (Interviewee 6, Interviewee 7) and fines (Interviewee 5, Interviewee 8). Interviewee 5 made the differentiation that only companies writing false reviews should be punished. Interviewee 3 and Interviewee 4 thought that the matter should be settled in court. Interviewee 7 on

the other hand thought that the law should only get involved when the insults are personally targeted. These imagined and proposed punishments are compiled in table 10.

False reviews were equated to defamation (Interviewee 4), fraud (Interviewee 8), and hindering one's livelihood (Interviewee 2, Interviewee 3). Also, involving a third party to serve as "*watchdog*" or as public oversight of some kind was suggested by Interviewee 3 and Interviewee 4. Interestingly the respondents who felt false reviews were akin to fraud etc., also thought of more serious punishments.

Table 10 Imagined punishments for false reviews

Respondent	Bans on platforms	Fines	Legal action
Interviewee 1			
Interviewee 2			
Interviewee 3			X
Interviewee 4			X
Interviewee 5		X	
Interviewee 6	X		
Interviewee 7	X		X*
Interviewee 8		X	

The dilemma of determining whether a review was deliberately false and not just negative etc. was also brought up by the interviewees Interviewee 3 and Interviewee 6 which in turn made the punitive actions harder to determine. Interviewee 6 thought that punishments stronger than bans on the platform were problematic, as the false reviews can be difficult to prove. Interviewee 3 and Interviewee 4 on the other hand thought that it's problematic for platforms to have such strong say on what can be said and what cannot. "*They are platforms, they can't really limit what can be said. They cannot take the Putin route so to say. There needs to be a freedom of speech.*" (Interviewee 4).

> *I think it's also wrong that some company like Google or Facebook has the power to stop you from expressing your opinion. If someone is wrong on something or does something wrong, it should rather be handled through some judiciary than some bot in Facebook. (Interviewee 3)*

It seems that some regulation is indeed needed on the field of online reviews and the platforms as many thought that the platforms had too much power for example. It was also previously mentioned by Interviewee 5 and Interviewee 8 that the cooperation with the platforms can be difficult for example because finding the truth about a review's legitimacy can be difficult.

Third parties were suggested to tackle problems related to false reviews, which could work. However due to the sheer number of online reviews this is

likely impossible without AI, which again circles back to the dilemma of a computer or an algorithm deciding what is true and what is not.

When asked what consequences a company would face if caught buying false reviews, the responses were similar among the respondents. Interviewee 1, Interviewee 2, Interviewee 4, Interviewee 6, Interviewee 7 thought that a company getting caught from buying reviews would lose its credibility. However, Interviewee 7 thought that getting caught from buying positive reviews had minimal consequences, while getting caught from negative ones was much more severe.

> *If I noticed or heard that someone had bought reviews (…). It instantly strikes me that, what's wrong with that place, why did they feel the need to do it. The credibility of the restaurant kind of disappears at that point. (Interviewee 2)*

> *I think that when we have a sauna event with the other entrepreneurs. They [the buyer of reviews] might get some snarky comments or laughed at, but I think that's about it. I don't think anyone is going to get burned at the stake for it. (Interviewee 7)*

Closely related to this, Interviewee 3, Interviewee 6 and Interviewee 8 though that getting caught would deter customers. Additionally, Interviewee 8 thought that "*(…) if you get caught as a company, the audience for the news about you getting caught would be many times bigger than that of the reviews.*" (Interviewee 8). Consequences for getting caught for the writer were elaborated on by Interviewee 6 and Interviewee 8, who thought that they were minimal to none.

The table 11 presents the remedial actions that were identified among the respondents and shows which respondents employed or would employ which methods. The figure 4 on the other hand attempts to answer the first research question: **"**what actions are companies employing to prevent damages from fake internet reviews and why?".

Table 11 Remedial actions that would be employed in case of false reviews

Interviewee	Replying	Bouncing back	Platforms	Removing/Deleting the review/comment
Interviewee 1	X	X	-	X
Interviewee 2	-	X	X	-
Interviewee 3	X	-	-	X*
Interviewee 4	-	X	X	-
Interviewee 5	-	X	X	X*
Interviewee 6	X	-	X	-
Interviewee 7	-	X	-	X
Interviewee 8	X	-	X	-

To answer the first research question: "what actions are companies employing to prevent damages from fake internet reviews?", the actions of the companies related to false reviews were inquired. The results were loosely categorized in to the four categories of the security action cycle: deterrent actions, detective actions, preventative actions, and remedial actions. The different actions were mainly unintentional or imagined as none of the interviewed companies had been the target of a more serious case of false reviews.

For deterrent actions, only the pre-built reputation was brought up, which was seen to act as a shield of sorts by making the false reviews to not have an effect on the overall rating.

In the category of preventative actions, some actions are blurred between the company and the platforms. For example, using platforms with high posting costs and having them pre-screen the reviews. However, one participant mentioned that they interview the hotel guests after their stay, which he thought that would prevent poor reviews as the guest had a chance to provide feedback on the spot. As the false reviews were seen to be the result of individuals who for some reason were upset with the company or owner, "doing a good job" daily was seen as a way of not giving the customer a reason to get angry in the first place. This would also feed into the deterrence feedback loop by strengthening their reputation.

Actions in the detection category were limited to meetings regarding the reviews in general. Also, many respondents would have a meeting in the case of multiple suspicious reviews. Factors affecting the detection were also discussed in order to gain more insight into this.

The remedial actions on the other hand were the most common as it is difficult for a small company to act before the false review has been detected. The most popular choice of action was in essence to do nothing as the individual false review would not have any effect. However, in the case of a more serious slander campaign, most companies would attempt to contact the platform, on which the false reviews were located on but in general most did not have high hopes on this achieving anything. Instead, by continuing the day-to-day operations and not reacting to the campaign, many thought that the rating would eventually return to normal.

Responding to the reviews to correct factual mistakes for example was seen as somewhat risky. Many felt that by responding you would get provoked, thus validating the individual false reviewer's actions. The idea behind not responding was to "not to feed the troll" so to say. However, in the case of a more serious campaign, some would make a public statement and in general, responding to the obvious campaign was seen as a more viable action. Closely related to this, moderating the comments by deleting and removing them was seen as double-edged sword, which could backfire if used too much. Also as mentioned earlier, the individual false review was seen to not have much effect so leaving it be, was the safer route. Figure 4 compiles this answer to the first research question.

Figure 4 Actions employed by the companies to mitigate false review damages

5.6 Opinions on purchasing reviews for themselves

The secondary research question in this thesis looks at factors deterring companies from purchasing reviews online either to boost their own ratings or to slander their competition. In this section the interviewees' opinions on the matter are discussed.

The idea of buying any false reviews was met with negative opinions in general. Some respondents however thought that buying positive reviews to boost their ratings was not as bad as buying negative reviews to slander the competition. The main reasons why the respondents would not buy these reviews were personal ethics and getting caught. When the positive false reviews were discussed, the idea of self-deception was brought up by many.

The attitudes towards the sellers of false reviews were also investigated to see how they might affect the desire to purchase reviews. In general, the sellers were seen as shady and unreliable. Interviewee 4, Interviewee 6, and Interviewee 7 were not aware of the sellers existing, while the others had varying levels of knowledge regarding these services. Most respondents saw the sellers as unreliable solely due to the unethical nature of the service and the likelihood of them scamming their customers high. Interviewee 1, Interviewee 3, Interviewee 6, and Interviewee 7 believed to varying degrees that they would get what they ordered from these sellers. Interviewee 3 however states that the chances of getting scammed are higher compared to other services.

5.6.1 Positive reviews to boost one's rating online

While buying negative reviews for competition was condemned by all respondents, buying positive reviews to boost one's own company was seen in a less negative light. Two respondents, Interviewee 1 and Interviewee 7 were the only ones to clearly state that boosting is fair game to an extent, however Interviewee 1 later stated that he would not buy them. Also, Interviewee 4 seemed indifferent by implying that it's part of the game so to say, however they later condemned it as false advertising. Interviewee 5 also admitted to having given positive reviews for friends' restaurants despite not having ever visited them, she didn't equate these reviews as false. *"(…) I might have given [positive reviews] to restaurants of people I know or our own but that is probably a different thing."* (Interviewee 5).

Interviewee 1 for example though that a company that does poorly will get caught very quickly when purchasing reviews as it is very obvious. However, they also felt that boosting your own rating while doing a good job is not a big deal. Interviewee 1 would not however purchase reviews for themselves as they already had a good reputation and had an ongoing aggressive marketing campaign.

> *You can highlight the good work but then if you produce garbage and give yourself good ratings, no one is going to believe it. (…)*

if you do good work and get good ratings, you can put a few of them
there yourself as well. (Interviewee 1)

Interviewee 7 also did not see buying positive reviews as a very bad thing and equated them with marketing. When asked about the reliability of the sellers of reviews he responded by saying: *"I think it's an amazing invention, I'm not against it. It's a way of marketing just like anything else. They [sellers] undoubtedly have a good system."* (Interviewee 7). At most Interviewee 7 thought that repercussions from getting caught from positively boosting their restaurants rating were limited to snarky comments from colleagues.

Apart from Interviewee 1 and Interviewee 7, the other interviewees saw buying positive reviews to boost one's reputation as self-deception and ethically wrong. These respondents specifically denounced buying reviews. Interviewee 1, Interviewee 2, Interviewee 5, and Interviewee 8 thought that buying false positive reviews is just self-deception. Buying reviews was deemed ethically wrong by Interviewee 3, Interviewee 4, Interviewee 6, and Interviewee 8

Interestingly Interviewee 1 and Interviewee 7 were both co-owners of their respective, smaller companies. The other smaller companies Interviewee 4 and Interviewee 6 were not owners, but rather spouses of the owners who were working in the company. While the sample size regarding company owners is limited, it could be deduced that company owners look at buying false reviews to boost their own rating in a more positive light as they are the ones is charge and answer to no-one. They also have the most to lose if the company does poorly and the most to gain if it does well, thus looking at false positive reviews from a slightly different angle. Compared to the employees of the companies on the larger end of the sample in terms of size, who were for example restaurant managers etc. who are not in charge of the company have less to gain or lose if the company suddenly does much better or worse. Also, the entrepreneurial mindset of Interviewee 1 and Interviewee 7 might cause a more "ruthless" view on using false positive reviews as a tool to market one's restaurant considering they did not mention any ethical problems among other factors.

It is to be noted however that Interviewee 1, Interviewee 4, and Interviewee 7 were not knowledgeable about false review sellers in general, which again reinforces the idea that they are not very common in Finland, at least regarding smaller companies. But Interviewee 1, Interviewee 4, and Interviewee 7 also might not be aware of the issues related to false reviews due to this and as they equated reviews to online marketing. During the interview, Interviewee 1 and Interviewee 7 both talked about online and other marketing they have done, which also could indicate that they feel that false positive reviews and marketing are closely connected.

5.6.2 Negative reviews to slander competition

Most interviewees categorically denounced buying false reviews as a whole and did not elaborate specific differences specifically to negative false reviews as they were already asked about the subject in general. All interviewees saw buying

negative false reviews in a more negative light compared to the positive ones, for example Interviewee 8 saw it as very shameless and ethically more wrong than buying positive reviews.

Negative false reviews were seen in a negative light by Interviewee 1 and Interviewee 7 as well, who did not see buying positive reviews as a very bad thing. Overall, the interviewees did not elaborate as much on negative false reviews as they did on positives, likely as the same deterrents might apply.

The general reaction to false negative reviews was much more stark and more ethically wrong than the positive false reviews. It was also imagined by some to be much more transparent and easier to get caught from. For example, Interviewee 7 though that by being in a smallish city, the chances of getting caught from buying negative reviews for competition are high and implied that these campaigns are expensive. *"[Interviewee 7 city] is such a small city that no company is going to have the guts to buy any slander campaigns. I don't think anyone has the money for that either."* (Interviewee 7). When asked to elaborate, Interviewee 7 thought that it would be obvious who had done it, because the community is close knit and everyone in the industry basically knows each other compared to the big cities. Interviewee 2 also suspected that a good reputation among the competitors prevents negative review campaigns.

I see them as reputable so I don't think it could get dirty. Of course, it is not impossible, maybe I'm just a naïve fool. (…) even though we have competition, we want to do our own thing and maybe I see it in a way that the others respect us back in a similar fashion.
(Interviewee 2)

5.6.3 Deterrents affecting the interviewees

In general, the reasons stated by the interviewees when asked why they would not purchase reviews were similar. The two main reasons were ethics and getting caught. The only respondent who specifically stated that their reputation was already good enough was Interviewee 1, but this was implied by others as well during the interviews. Also as mentioned earlier, getting caught had negative consequences such as losing credibility and deterring customers.

False reviews were specifically stated to being ethically wrong and this was also the biggest reason why the respondents would not buy them. This answer was given by Interviewee 3, Interviewee 4, Interviewee 5, Interviewee 6, and Interviewee 7. However, in Interviewee 7 case, only negative reviews were implied to being ethically wrong. Interviewee 1 also was on the fence whether buying positive false reviews was ethically very wrong, as he equated them to marketing to a degree.

The other main deterrent was the risk of getting caught and the subsequent consequences, which were discussed in more depth earlier. Interviewee 1, Interviewee 2, and Interviewee 7 thought that it would be obvious when someone had bought false reviews and people would notice it quickly. Interviewee 1 however specified that the risk of getting caught from positive reviews existed only if the

bought ratings did not match the perceived quality of the establishment. Reliance on repeat purchase as a deterrent factor was also mentioned by Interviewee 5 and implied by Interviewee 2.

As the interviewed companies already had a good reputation in their own opinion, backed up by ratings from google, shown in Table 4, it is likely that the risk versus reward is too big in this case. Most of the companies also had plenty of reviews on their google page, with only Interviewee 4 and Interviewee 6 in the double digits, making buying false reviews less lucrative. Why would they try to increase their google rating by a few decimals when the rating was already good, and the possible backlash was seen as severe even though considering Interviewee 1, who had extensive marketing experience, mentioned that buying reviews when the rating is good, is less visible which in turn makes the risk of getting caught smaller. Also, in case of Interviewee 2, whose company had several restaurants across Finland, the good reputation gained in those other locations is likely to make purchasing positive reviews redundant.

A factor related to this is the thought that the current market area is too small for a false review campaign to be successful and undetected as Interviewee 5, Interviewee 6, Interviewee 7, and Interviewee 8 mentioned during the interviews. Interviewee 5 and Interviewee 8 mentioned Finland in general as a market to being too small, and Interviewee 6 and Interviewee 7 their respective market areas as too small as well. This, and the fact that the fields of restaurants and hotels seem too well-connected as mentioned by Interviewee 5 and Interviewee 7, play it to the idea of low competitive environment, which in turn makes false reviews less lucrative. Meaning there is no need to buy false reviews and the risk of getting caught is considerable.

Closely related to the company reputation, as most of the respondents were in the restaurant industry it is likely that their customer segment is the local population who already know the restaurants via experience or word-of-mouth from friends and family. This would make any attempt to boost their ratings have little effect in terms of popularity or increased business. This effect is likely smaller on the travel industry Interviewee 5 and Interviewee 8 in this study as locals likely do not use hotels in their own city.

Reliance on repeat customers was mentioned by Interviewee 5 and implied by Interviewee 2. As most of the interviewed companies were restaurants, repeating customers are vital for their business and likely also serve as deterrent factors for them in this regard, despite them not mentioning it directly.

A passing mention was made by Interviewee 7 who thought that anyone has the money to sustain a negative false review campaign against an established restaurant. Likely referring to the fact that restaurants would bounce back after a while.

The interviewees were also asked about their opinion on the sellers of false reviews. While deterrence was not specifically discussed in this regard, many saw the suspiciousness of the craft to be off putting.

Roughly half of the respondents saw the sellers as unreliable solely due to the unethical nature of the service and the likelihood of them scamming their

customers high. Interviewee 2 for example said that *"(…) if you are offering a kind of scam service, why would the firm itself be legitimate."* (Interviewee 2). When asked about their opinions on the sellers of false reviews, Interviewee 4, Interviewee 6, and Interviewee 7 were not aware of the sellers existing, while the others had varying levels of knowledge regarding these services.

Interviewee 1, Interviewee 3, and Interviewee 7 believed to varying degrees that they would get what they ordered from these sellers. Interviewee 3 however states that the chances of getting scammed are higher compared to other services. Once again, Interviewee 1 and Interviewee 7 are more lenient towards the sellers compared to the others likely due to them being entrepreneurs, as discussed earlier. Table 12 compiles the interviewees views on the sellers.

Knowledge and opinion about the sellers:

- Interviewee 1: Knew about the sellers existing and had checked prices out of curiosity in the past. Thought that they are not sold by Finns, but implied people would be more interested in them if the seller was Finnish.
- Interviewee 2: Did not see the sellers as reliable. Instead saw them as shady and akin to Ponzi schemes and other get-rich-quick scams. Was skeptical whether you would get what you order.
- Interviewee 3: Saw the sellers as just another business among others, maybe a bit less ethical. Believed that you would get what you order but the chance of getting scammed is higher.
- Interviewee 4: Was not aware of sellers existing. Saw them as suspicious.
- Interviewee 5: Did not see the sellers as reliable due to the unethical nature of the service.
- Interviewee 6: Was not aware of the sellers. Saw the practice of selling false reviews as "shocking". Was certain that they would not receive what they order.
- Interviewee 7: Was not aware of the sellers but believed that they would get what they order. Equated the practice of selling false reviews to marketing.
- Interviewee 8: Saw the sellers as shady, as the service is not "based on anything". Expected that they would not receive what they order because the service is already based on a hoax making it likely that they will scam their customers as well.

Table 12 Opinions on the sellers of reviews

Interviewee	Was aware of sellers (Y/N)	Would get what they ordered (Y/N)	Are sellers seen as shady or unethical
Interviewee 1	Y	Y	-
Interviewee 2	Y	N	Y
Interviewee 3	Y	Y*	Y
Interviewee 4	N	-	Y

Interviewee 5	Y	N	Y
Interviewee 6	N	N	Y
Interviewee 7	N	Y	N
Interviewee 8	Y	N	Y

All in all, personal deterrents affecting the purchase of false reviews are mainly rooted in the ethics of the respondents as well as in the fear of repercussions. These repercussions are imagined as much worse than the potential gain that they would receive from false positive reviews, especially as the companies in question already had good reputations.

The deterrents were much stiffer in the case of negative false reviews. From their own viewpoint, the respondents would avoid a company that is known for committing review fraud and thought that others would as well. While the lack of stiff competition was not mentioned as a deterrent, it was implied that the word would get around quick due to closely connected fields. The sellers themselves seemed to deter the interviewees from buying to a degree as they were seen as unreliable and shady by most.

To answer the second research question: the deterrents for personally purchasing false reviews in order to boost one's own company rating or to slander the competition were mainly applicable in both cases, however some only applied to one or the other. Deterrents that were applicable solely to purchasing positive false reviews were the reliance on repeat purchase, ie. wanting the customer to return and already having a good enough reputation making the purchase redundant. As for factors solely related to the negative false reviews, many felt that the companies are well connected and respective of each other thus making the threshold to negatively false review much higher. Also, the cost of sustaining such campaign was brought up by one respondent. As for the shared deterrents, ethics were seen as the biggest deterrent and the repercussions from getting caught a close second. These repercussions included loss of face and loss of customers for example. It was also mentioned by several respondents that the market would be too small for this type of action, which is likely somewhat related to the companies being closely connected. This in turn makes the chances of getting caught much higher because the companies are connected in such a small market. Lastly, the sellers of false reviews in general were seen as suspicious with many respondents doubting whether they would receive the good reviews they had paid for. The reasoning for this was that a company that deals with such shady practices is likely to be shady in itself. Figure 5 compiles the mentioned deterrents and sorts them by their relation to either positive or negative false reviews or both, thus answering the second research question: "what factors deter a company from colluding with fake reviews?".
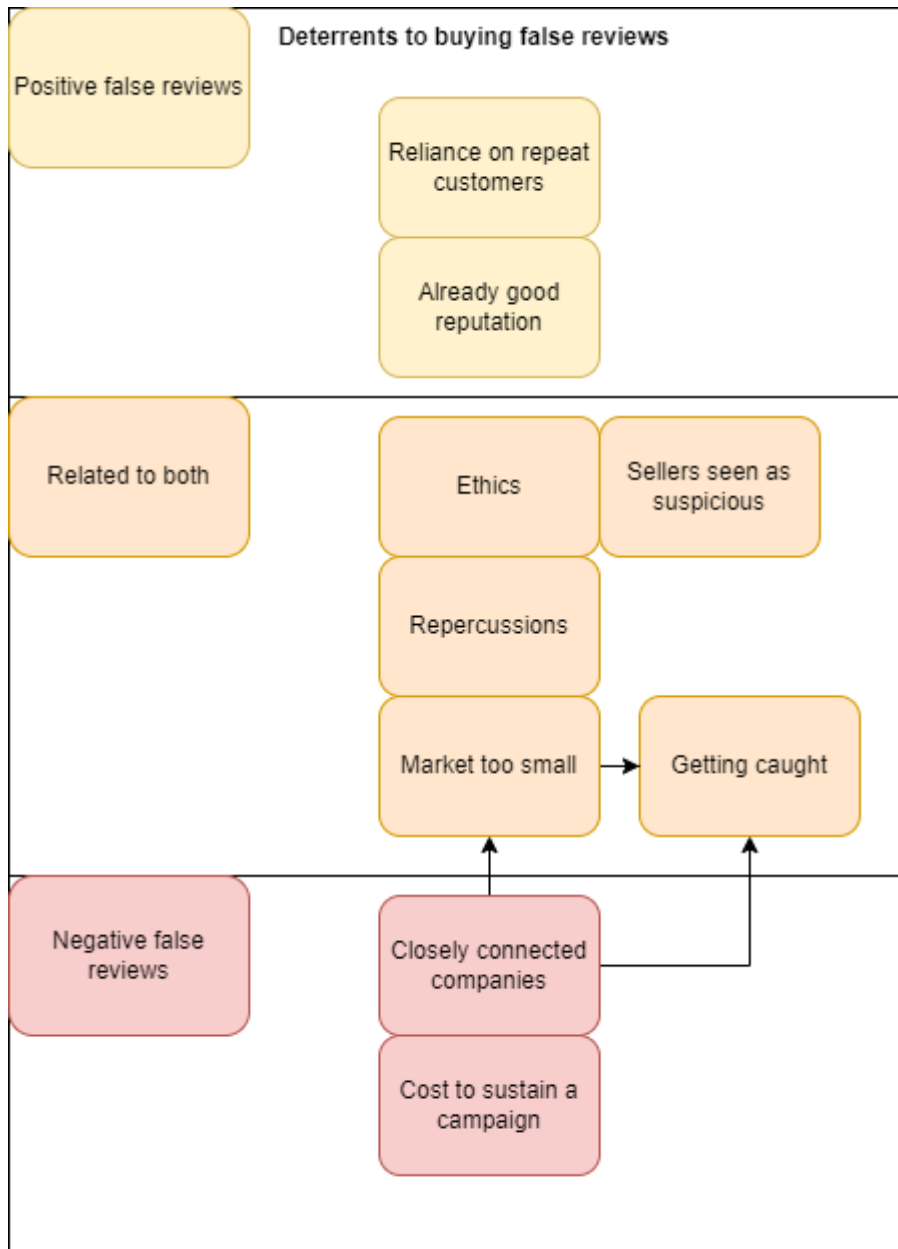
Figure 5 What factors caused the interviewees to not collude with false reviews

6   DISCUSSION

This research set to clarify what actions, if any, are employed by small finnish companies to prevent damages from false reviews. Also, the opinions on buying these services were on the forefront. To guide the research, two research questions were conceived: "what actions are companies employing to prevent damages from fake internet reviews" (RQ1) and "what factors deter a company from colluding with fake reviews" (RQ2). The research was structure around the security action cycle by Straub and Welke, (1998) which included four stages related to computer misuse, which in this case were the false reviews. These stages were: deterrence, prevention, detection, and remedies. This chapter presents the theoretical and practical implications of this study while also proposing future research directions. The limitations of the study are also discussed.

6.1   Theoretical implications

In general, the acquired results agree with previous literature with some exceptions and additionally some factors seemed to not have been considered before. For example, a factor presented in the deterrence chapter that was seen to deter false reviewers, small market size/closely connected companies, was not brought up in the studied literature. However, a more likely reason that is related to closely connected companies is the lack of cutthroat competition, as restaurants are more likely to receive false reviews when there is a lot of competition according to Luca and Zervas, (2016). It is arguable whether this factor is within the scope of the first research question as the lack of fierce competition is likely not the result of the companies' actions. On the other hand, this could be seen as connecting with the other companies to promote cooperation etc.

Many of the interviewees thought that by doing a good job and consequently having a good rating beforehand, the false reviews would not have such a large effect in the first place. This is somewhat reminiscent of the enhancement strategies by Lappas et al., (2016). The general idea behind the enhancement

strategy is to improve the existing features and services to gain better reviews and prevent the poor ones from being written. This strategy was however not consciously implemented but was a part of their day-to-day business.

Largely prevention was seen to be the responsibility of the platforms, as many high posting costs and other preventative measures can realistically only be employed by the platforms. High posting costs were present in the answers through having to use one's real identity (a proposed solution to false reviews) to post reviews or having to have purchased the product or service. High posting costs were seen as a working method by Ott et al., (2012). However, some concerns were raised by the interviewees about privacy regarding using one's real identity, but this is outside of the scope of this study.

The actions related to detecting the reviews were supported by the literature. For example, most factors that were used by the interviewees to determine review truthfulness were also confirmed by Filieri, (2016) with at least one method of detection mentioned from each of the four categories: source trustworthiness, message trustworthiness, review valence, and patterns in reviews. For example, reviewers were seen as suspicious if they did not have a real name or a profile picture, which is also in accordance with Munzel, (2016).

Many respondents felt that the regular customers could see past the false reviews. This trust can however be misplaced as humans in general are poor at detecting false reviews (Ott et al., 2011). Also, regular customers are likely to not read reviews about a restaurant they are familiar with.

While the enhancement strategies by Lappas et al., (2016) were seen to work, the confrontational strategy by Lappas et al., (2016), which in essence, was to respond and challenge the false reviews received more doubtful answers. Even though Wang and Chaudhry, (2018) found them to have a positive effect on the subsequent readers of the comment, the responding was seen as too risky. It is likely that the difference lies in the type of perceived false review attack (individuals vs. a campaign) and their effects. A possible, underlying reason for this could be related to the fact that weaker and less known brands responses are less effective at this, according to Ullrich and Brunner, (2015). For example, as Interviewee 4 thought, replying would just be their word against the reviewers.

Some respondents on the other hand had had outsiders defend the company in the reviews and online comments. Irrelevant whether this was orchestrated by the company as was the case with Interviewee 1 or not, the effect on the readers is larger than if the company itself replied (Bickart & Schindler, 2001), which was clearly the intention of Interviewee 1 for example.

Regarding the second research question, the pre-existing literature related deterrents related to buying false reviews for one's own need seems non-existent. In general, the results point at ethical reasons and other consequences being the main deterrents. However, a closer look at the consequences could be taken in the future to see what factors a major effect and which ones are more minor. This is a potential new research topic that could yield information on how to improve the platforms' resistance to false reviews as well as make them have less effect on the companies' utilizing reviews. As all the interviewed case companies had a

relatively good rating on google and relied on repeat purchase, the appeal of positive false reviews is smaller, as stated by Chen et al., (2019). This was also pointed out by Interviewee 5 in the interviews. However, any of the factors that would make the false reviews more appealing according to Chen et al., (2019) were not observed in the interviewees who were more lenient towards them. These factors related to making false reviews more appealing in the potential buyers' eyes were listed earlier in the buyers of false reviews chapter.

The findings presented in this study do fit the security action cycle in some regard. While in some cases, the boundary between the stages is fickle, as many actions could be interpreted to belong in another stage, depending on whose viewpoint is used for example. Despite this, most actions fit the cycle well. However, due to the target of the study being the companies utilizing these reviews, not much can be placed in the phases preceding the incident of a false review.

Most of the imagined or experienced false reviews that the interviewees reflected upon were based on two of the three categories for false reviews by Jindal and Liu, (2008):

1. untruthful opinions that aim to deliberately mislead with the intention to boost or defame a target,
2. and reviews on brands only, where the reviews are targeted at the brand instead of the product or service.

The third category, non-reviews, was not as prevalent in the interviews. The often-personal reasons for false reviews mentioned by many interviewees can be seen to belong in the reviews on brand category.

A factor that rose from the second research question was how the position of the interviewee affected their view on buying false reviews as the two entrepreneurs clearly saw buying reviews to boost ratings as less unethical than the others, who were not in the same position.

Also, while size of the company was initially thought to possibly have an effect. It turned out to not have much to do with either of the research questions. What seemed to however, was the field in which the companies existed, as the hotels were much more knowledgeable. These two factors could be researched further in more depth in the future.

As many though that Finland as a market was too small for a false review campaign to work properly, it could be worth researching the subject whether this is true and if so, why. Also, considering the false reviewer was unanimously seen as an individual, another potential research direction could look in to why this is. Are the campaigns so subtly done or do they simply not exist in Finland?

6.2   Practical implications

As this thesis set out to chart the current state of matters regarding defenses and attitudes towards false reviews, many practical implications could be considered. Considering all the case companies had a good rating on google reviews, the

current actions can be considered sufficient and working. However as stated earlier, the absence of a more serious negative review campaign, whether observed or experienced, is likely to contribute to the current state of matters. In practice, these results could be used to bolster the defenses of companies like the sample against false reviews. For example, as discussed in the previous chapter, it was seen by the interviewees that a closely connected market would harbor less negative false reviews, which could be put in to practice by any company by connecting and working together with another company for example.

As the lack of knowledge about false reviews in general was somewhat present in the interviews, which could be seen as an opportunity spread awareness about the false reviews.

Also based on the literature and empirical results, different defensive mechanisms could be devised to counter disgruntled individuals and more serious campaigns separately, as factors that seem to deter individuals, may not always deter the professionals and vice versa as pointed out by several interviewees.

6.3   Limitations

This chapter presents limitations that were found regarding the empirical study and its results. Also, future research directions based on matters that seemed to have more to them are proposed.

This thesis had a sample of eight case companies from the service industry in Finland. The small market size of Finland may skew the results, making a similar study conducted in more populated countries or cities yield completely different results. Another factor preventing generalization is the culture, as even in similarly populated country, the ethical views of the population might cause significantly different results. These factors are likely more limiting when considering the second research question, whereas the countermeasures can be seen as more generalizable.

While the respondents were all working with their respective social medias, the elite bias was not present. On the contrary, many of the respondents had surprisingly lacking knowledge on the subject considering their positions. This potentially caused them to talk about Facebook comments that they had received when the talking point was about online reviews. This however might point that these two are similar in terms of effect. Also, as most of the respondents did not really have any experience with false reviews, the response strategies were based on hypothetical situations such as false review campaigns. Considering this, a more false-review-heavy environment could yield significantly different and richer results.

As this study only grazed the surface of false review countermeasures, a future study could focus more closely on these countermeasures and find out which ones actually work and which ones are just though to.

It is not likely that the interviewees withheld important information either, as some of the interviewees shared information that many companies would

likely not publicly announce or present. However, a factor that should be considered regarding the sample, is the selection process. As the interviewees were recruited online, many companies that would perhaps provide more controversial results might not respond to the interview requests. Also, companies with weaker reputation might also yield different kinds of results, which could be a potential future research opportunity.

7   CONCLUSION

This thesis set out to figure what actions are in place to prevent damages from false reviews in the interviewed Finnish service sector companies and why. Also, a secondary research question, what factors prevent these companies or interviewees from purchasing false reviews for their own needs was devised to guide the research.

Initially this thesis introduced the subject of internet reviews, in which their importance and influence were emphasized. In the first chapter of the literature review, topics closely related to online reviews, such as review platforms, review effect, review trustworthiness and review management were covered. By covering these topics, a general understanding of how, why, and where the online reviews influence their reader.

After covering the online reviews in general, can the fake online reviews be discussed. The second chapter of the literature review introduced this dark side of online reviews by building on what was covered earlier. The topic covered the different causes and effects as well as other features and factors related to false reviews, such as the buyers, sellers, and writers. Also, the malevolent versions of crowdsourcing: astroturfing and crowdturfing were covered as they are often closely related to the spread of false reviews.

The third chapter introduced the theory on which this thesis is based on. The security action cycle by Straub and Welke, (1998), was chosen as the theory, through which these research questions and the thesis as a whole are looked at. While not originally designed for false reviews, the security action cycle was proposed to combat systems risk, which in this case is manifested as fake online reviews. This theory was used to guide the interviews by asking questions based on the four different stages of the cycle and when analyzing the results.

The empirical part of the thesis was conducted through semi-structured interviews, as research on the subject is still slim and relatively unexplored. Eight participants from selected eight case companies were interviewed from the Finnish service sector. These companies were mainly restaurants/bars but also two hotels and one goldsmith's boutique were included. Some interviewees also were responsible for multiple companies, and some were part of a larger concern. In

general, these companies could be considered small or very small in terms of size based on publicly available knowledge of them. Initially when choosing the sample, the companies' sizes were taken into consideration to see whether this had an effect but there was little that could be attributed to their size alone based on this study. What seemed to be more influential was the industry and interviewees position for example.

As mentioned earlier, the recruitment method of contacting the participants online may skew the results as the companies with tendencies to these unethical practices are less likely to partake in the interviews.

In general, the knowledge about false reviews among the sample was lacking and the preparation and plans even more so. The current toolset is enough to deal with the individual false reviewer but in the case of a more severe campaign many would struggle. This is the case likely due to negative false review campaigns not having been observed and false reviews being relatively rare due to many different factors, such as small market size.

The current actions that the case companies are taking are quite slim and more so, when talking about conscious effort. Most companies did not really have any strategy if a false review campaign was to happen immediately and hoped that by doing a good job in the first place, the customer would not have a reason to get upset. The likely reason for the lack of plan is the lack of serious cases as mentioned earlier. However, the imagined responses in a more serious situation were related to responding to the reviews, contacting the platforms and public statements.

As for the opinions on buying false reviews for their own gain, most felt that positive false reviews were bad, while everyone agreed that negative false reviews were bad. The main deterrents were related to ethics and the fear of consequences if they were to get caught.

In general, the results agree with previous literature, even though prior research related to either of the research questions specifically was not found. While nothing was found that would contradict the previous literature regarding online reviews, some factors seemed unexplored and could only be considered to be partly supported by literature, at best.

Some limiting factors regarding the generalization of the results were found. For example, the size of the interviewed companies being small, their service sector and interviewee position likely affect the results to a degree. Even the culture present in Finland can cause the results to not be applicable in countries where false reviews are more accepted and/or common.

REFERENCES

Ahluwalia, R. (2002). How Prevalent Is the Negativity Effect in Consumer Environments?. Journal Of Consumer Research, 29(2), 270-279. doi: 10.1086/341576

Ahmad, W., & Sun, J. (2018). Modeling consumer distrust of online hotel reviews. International Journal Of Hospitality Management, 71, 77-90. doi: 10.1016/j.ijhm.2017.12.005

Amazon.com Help: Anti-Manipulation Policy for Customer Reviews. (2021). Retrieved 30 November 2021, from https://www.amazon.com/gp/help/customer/display.html?nodeId=201996120

Amazon.com Help: Community Guidelines. (2021). Retrieved 1 December 2021, from https://www.amazon.com/gp/help/customer/display.html?nodeId=GLHXEX85MENUE4XF

Anderson, A., Harris, A., & Miller, J. (1983). Models of Deterrence Theory. SOCIAL SCIENCE RESEARCH, 12, 236-262.

App Store Review Guidelines - Apple Developer. (2021). Retrieved 14 October 2021, from https://developer.apple.com/app-store/review/guidelines/

Baek, H., Ahn, J., & Choi, Y. (2012). Helpfulness of Online Consumer Reviews: Readers' Objectives and Review Cues. International Journal Of Electronic Commerce, 17(2), 99-126. doi: 10.2753/jec1086-4415170204

Banerjee, S., & Chua, A. (2014). A study of manipulative and authentic negative reviews. Proceedings Of The 8Th International Conference On Ubiquitous Information Management And Communication - ICUIMC '14. doi: 10.1145/2557977.2557984

Bickart, B., & Schindler, R. (2001). Internet forums as influential sources of consumer information. Journal Of Interactive Marketing, 15(3), 31-40. doi: 10.1002/dir.1014

Byers, J., Mitzenmacher, M., & Zervas, G. (2012). The groupon effect on yelp ratings. Proceedings Of The 13Th ACM Conference On Electronic Commerce - EC '12. doi: 10.1145/2229012.2229034

Chen, L., Li, W., Chen, H., & Geng, S. (2019). Detection of Fake Reviews: Analysis of Sellers' Manipulation Behavior. Sustainability, 11(17), 4802. doi: 10.3390/su11174802

Doh, S., & Hwang, J. (2009). How Consumers Evaluate eWOM (Electronic Word-of-Mouth) Messages. Cyberpsychology & Behavior, 12(2), 193-197. doi: 10.1089/cpb.2008.0109

East, R., Hammond, K., & Lomax, W. (2008). Measuring the impact of positive and negative word of mouth on brand purchase probability. International Journal Of Research In Marketing, 25(3), 215-224. doi: 10.1016/j.ijresmar.2008.04.001

Estellés-Arolas, E., & González-Ladrón-de-Guevara, F. (2012). Towards an integrated crowdsourcing definition. Journal Of Information Science, 38(2), 189-200. doi: 10.1177/0165551512437638

Filieri, R. (2016). What makes an online consumer review trustworthy?. Annals Of Tourism Research, 58, 46-64. doi: 10.1016/j.annals.2015.12.019

Fontana, A. & Frey, J.H. (2000). The interview. From structured questions to negotiated text. In Y.S.Lincoln & N.K. Denzin (Eds.), Handbook of qualitative research (2nd ed., pp. 645–672). Thousand Oaks, CA: Sage.

He, S., Hollenbeck, B., & Proserpio, D. (2020). The market for fake reviews. Available at SSRN.

Héder, M. (2018). A black market for upvotes and likes. Információs Társadalom, 19(4), 18. doi: 10.22503/inftars.xix.2019.4.2

Hennig-Thurau, T., Gwinner, K., Walsh, G., & Gremler, D. (2004). Electronic word-of-mouth via consumer-opinion platforms: What motivates consumers to articulate themselves on the Internet?. Journal Of Interactive Marketing, 18(1), 38-52. doi: 10.1002/dir.10073

Hu, N., Pavlou, P., & Zhang, J. (2006). Can online reviews reveal a product's true quality?. Proceedings Of The 7Th ACM Conference On Electronic Commerce - EC '06. doi: 10.1145/1134707.1134743

Ipeirotis, P. (2010). Analyzing the Amazon Mechanical Turk marketplace. XRDS: Crossroads, The ACM Magazine For Students, 17(2), 16-21. doi: 10.1145/1869086.1869094

Ipeirotis, P. G. Mechanical turk: Now with 40.92% spam. Behind Enemy Lines blog, December 2010.

Jin Ma, Y. and Lee, H.-H. (2014), "Consumer responses toward online review manipulation", Journal of Research in Interactive Marketing, Vol. 8 No. 3, pp. 224-244. https://doi.org/10.1108/JRIM-04-2013-0022

Jindal, N., & Liu, B. (2008). Opinion spam and analysis. Proceedings Of The International Conference On Web Search And Web Data Mining - WSDM '08. doi: 10.1145/1341531.1341560

Kauppila, S., & Soliman, W. (2022). "This Is the Future of Advertising!" Or Is It? New Insights into the Justifiability of Deceptive Crowdwork in Cyberspace. Proceedings Of The Annual Hawaii International Conference On System Sciences. doi: 10.24251/hicss.2022.350

Kempf, C. (2018). Write Fake Reviews, Do Real Time. Retrieved 12 October 2021, from https://www.pcma.org/write-fake-reviews-do-real-time/

Kim, J., Kang, J., Shin, S. and Myaeng, S., 2021. Can You Distinguish Truthful from Fake Reviews? User Analysis and Assistance Tool for Fake Review Detection. Proceedings of the First Workshop on Bridging Human–Computer Interaction and Natural Language Processing, pp.53–59.

Kim, S., Maslowska, E., & Malthouse, E. (2017). Understanding the effects of different review features on purchase probability. International Journal Of Advertising, 37(1), 29-53. doi: 10.1080/02650487.2017.1340928

Lackermair, G., Kailer, D., & Kanmaz, K. (2013). Importance of Online Product Reviews from a Consumer's Perspective. Advances In Economics And Business, 1(1), 1-5. doi: 10.13189/aeb.2013.010101

Lappas, T., Sabnis, G., & Valkanas, G. (2016). The Impact of Fake Reviews on Online Visibility: A Vulnerability Assessment of the Hotel Industry. Information Systems Research, 27(4), 940-961. doi: 10.1287/isre.2016.0674

Lee, K., Webb, S. and Ge, H., 2015. Characterizing and automatically detecting crowdturfing in Fiverr and Twitter. Social Network Analysis and Mining, 5(1).

Lembcke, T.B., Trang, S., Plics, P., Masuch, K., Hengstler, S., Pamuk, M.: Fostering information security compliance: comparing the predictive power of social learning theory and deterrence theory. In: 25th Americans Conference on Information Systems AMCIS 2019, no. Bandura 1977, pp. 1–10 (2019)

Li, H., Fei, G., Wang, S., Liu, B., Shao, W., Mukherjee, A., & Shao, J. (2017). Bimodal Distribution and Co-Bursting in Review Spam Detection. Proceedings Of The 26Th International Conference On World Wide Web. doi: 10.1145/3038912.3052582

Li X, Hitt LM (2008) Self-selection and information role of online product reviews. Inform. Systems Res. 19(4):456–474

Lu, S., Wu, J., & Tseng, S. (2018). How Online Reviews Become Helpful: A Dynamic Perspective. Journal Of Interactive Marketing, 44, 17-28. doi: 10.1016/j.intmar.2018.05.005

Luo, H. (2021). A Review of Research on Identification of False Reviews in E-Commerce. J. Management And Humanity Research, Vol. 3, 2020, 9-15.

Luca M (2011) Reviews, reputation, and revenue: The case of Yelp.com. HBS Working Paper 12-016, Harvard Business School, Cambridge, MA.

Luca, M., & Zervas, G. (2016). Fake It Till You Make It: Reputation, Competition, and Yelp Review Fraud. Management Science, 62(12), 3412-3427. doi: 10.1287/mnsc.2015.2304

Mathews Hunt, K. (2015). Gaming the system: Fake online reviews v. consumer law. Computer Law & Security Review, 31(1), 3-25. doi: 10.1016/j.clsr.2014.11.003

Mathwick, C., & Mosteller, J. (2016). Online Reviewer Engagement. Journal Of Service Research, 20(2), 204-218. doi: 10.1177/1094670516682088

Martens, D., & Maalej, W. (2019). Towards understanding and detecting fake reviews in app stores. Empirical Software Engineering, 24(6), 3316-3355. doi: 10.1007/s10664-019-09706-9

Maslowska, E., Malthouse, E., & Bernritter, S. (2016). Too good to be true: the role of online reviews' features in probability to buy. International Journal Of Advertising, 36(1), 142-163. doi: 10.1080/02650487.2016.1195622

Motoyama, Y., & Usher, K. (2020). Restaurant Reviews and Neighborhood Effects. Papers In Applied Geography, 6(4), 386-401. doi: 10.1080/23754931.2020.1791942

Mukherjee, A., Venkataraman, V., Liu, B., & Glance, N. (2013). What Yelp Fake Review Filter Might Be Doing?. Proceedings of the International AAAI Conference on Web and Social Media, 7(1). Retrieved from https://ojs.aaai.org/index.php/ICWSM/article/view/14389

Munzel, A., 2016. Assisting consumers in detecting fake reviews: The role of identity information disclosure and consensus. Journal of Retailing and Consumer Services, 32, pp.96-108.

Murphy, R. (2016). Local Consumer Review Survey: How Customer Reviews Affect Behavior. Retrieved 28 April 2021, from https://www.brightlocal.com/research/local-consumer-review-survey/

Myers, M., & Newman, M. (2007). The qualitative interview in IS research: Examining the craft. Information And Organization, 17(1), 2-26. doi: 10.1016/j.infoandorg.2006.11.001

Nance,W. and Straub,D., "AN INVESTIGATION INTO THE USE AND USEFULNESS OF SECURITY SO TWARE IN DETECTING COMPUTER ABUSE" (1988). ICIS 1988 Proceedings. 36.

Narasimhan, C., & Turut, Ö. (2013). Differentiate or Imitate? The Role of Context-Dependent Preferences. Marketing Science, 32(3), 393-410. doi: 10.1287/mksc.2013.0776

O'Neil, T. (2015). Survey confirms the value of reviews, provides new insights. Retrieved from https://www.powerreviews.com/blog/survey-confirms-the-value-of-reviews/

Ott, M., Cardie, C., & Hancock, J. (2012 (Published). "Estimating the Prevalence of Deception in Online Review Communities,") Conference Estimating the Prevalence of Deception in Online Review Communities. 2187864: ACM. 201–210.

Ott, M., Choi, Y., Cardie, C., & Hancock, J. (2011). Finding Deceptive Opinion Spam by Any Stretch of the Imagination. Proceedings Of The 49Th Annual Meeting Of The Association For Computational Linguistics: Human Language Technologies,, 309–319.

Pritchett, D. (2021). A look at how we tackle fake and fraudulent contributed content [Blog]. Retrieved from https://blog.google/products/maps/google-maps-101-how-we-tackle-fake-and-fraudulent-contributed-content/

Review Fraud Consequences. Tripadvisor help center. (n.d.). Retrieved October 28, 2021, from https://www.tripadvisorsupport.com/en-US/hc/owner/articles/341.

Rinta-Kahila, Tapani and Soliman, Wael, (2017). "UNDERSTANDING CROWDTURFING: THE DIFFERENT ETHICAL LOGICS BEHIND THE CLANDESTINE INDUSTRY OF DECEPTION". In Proceedings of the 25th European Conference on Information Systems (ECIS), Guimarães, Portugal, June 5-10, 2017 (pp. 1934-1949). ISBN 978-989-20-7655-3 Research Papers.

Shin, H., Perdue, R., & Pandelaere, M. (2019). Managing Customer Reviews for Value Co-creation: An Empowerment Theory Perspective. Journal Of Travel Research, 59(5), 792-810. doi: 10.1177/0047287519867138

Siponen, M., Soliman, W., & Vance, A. (2022). Common Misunderstandings of Deterrence Theory in Information Systems Research and Future Research Directions. ACM SIGMIS Database: the DATABASE for Advances in Information Systems, 53(1), 25-60.Soliman, W., & Rinta-Kahila T., "UNETHICAL BUT NOT ILLEGAL: UNCOVERING THE PERSUASIVE MESSAGES LEVERAGED BY PROVIDERS OF THE "REAL" ONLINE SOCIAL IMPRESSIONS" (2018). Research Papers. 172.

Souppouris, A. (2021). Samsung fined $340,000 for faking online comments. Retrieved 12 October 2021, from

https://www.theverge.com/2013/10/24/5023658/samsung-fined-340000-for-posting-negative-htc-reviews

Sparks, B., Perkins, H., & Buckley, R. (2013). Online travel reviews as persuasive communication: The effects of content type, source, and certification logos on consumer behavior. Tourism Management, 39, 1-9. doi: 10.1016/j.tourman.2013.03.007

Straub, D., & Welke, R. (1998). Coping with Systems Risk: Security Planning Models for Management Decision Making. MIS Quarterly, 22(4), 441. doi: 10.2307/249551

Streitfeld, D. (2012). "The Best Book Reviews Money Can Buy. 〈http://www.nytimes.com/2012/08/26/business/book-reviewers-for-hire-meet-a-demand-for-online-raves.html?pagewanted=1&_r=2&partner=rss&emc=rss〉. Accessed February 11, 2013.

Sun, M. (2010). How Does the Variance of Product Ratings Matter?. SSRN Electronic Journal. doi: 10.2139/ssrn.1400173

Sutanto, M. A., & Aprianingsih, A. (2016). The Effect of Online Consumer Review toward Purchase Intention: A Study in Premium Cosmetic in Indonesia. In Journal International Conference on Ethics of Business, Economics, and Social Science.

Sykes, G., & Matza, D. (1957). Techniques of Neutralization: A Theory of Delinquency. American Sociological Review, 22(6), 664. doi: 10.2307/2089195

Tong, Y., Zhou, Z., Zeng, Y., Chen, L., & Shahabi, C. (2019). Spatial crowdsourcing: a survey. The VLDB Journal, 29(1), 217-250. doi: 10.1007/s00778-019-00568-7

Tunnuslukuja väestöstä muuttujina Alue, Tiedot ja Vuosi-PxWeb. (2022). Retrieved 3 May 2022, from https://statfin.stat.fi/PxWeb/pxweb/fi/StatFin/StatFin__vaerak/statfin_vaerak_pxt_11ra.px/table/tableViewLayout1/?loadedQueryId=45ba5f20-2525-44b7-96f0-0459758e1b5b&timeType=top&timeValue=1

Ullrich, S., & Brunner, C. (2015). Negative online consumer reviews: effects of different responses. Journal Of Product & Brand Management, 24(1), 66-77. doi: 10.1108/jpbm-05-2014-0611

Vargo, S., & Lusch, R. (2004). Evolving to a New Dominant Logic for Marketing. Journal Of Marketing, 68(1), 1-17. doi: 10.1509/jmkg.68.1.1.24036

Vermeulen, I., & Seegers, D. (2009). Tried and tested: The impact of online hotel reviews on consumer consideration. Tourism Management, 30(1), 123-127. doi: 10.1016/j.tourman.2008.04.008

Wang, G., Wilson, C., Zhao, X., Zhu, Y., Mohanlal, M., Zheng, H., Zhao, B.Y. (2012) Serf and turf: crowdturfing for fun and profit. In: Proceedings of the 21st international conference on World Wide Web (WWW), pp 679–688

Wang, Y., & Chaudhry, A. (2018). When and how Managers' Responses to Online Reviews Affect Subsequent Reviews. Journal Of Marketing Research, 55(2), 163-177. doi: 10.1509/jmr.15.0511

Weise, E. (2015). Retrieved 14 December 2021, from https://eu.usatoday.com/story/tech/2015/10/19/amazon-cracks-down-fake-reviews/74213892/

Willison, R., & Warkentin, M. (2013). Beyond deterrence: An expanded view of employee computer abuse. MIS Quarterly, 37(1), 1–20.

What does Tripadvisor do about unfair reviews? Tripadvisor . (n.d.). Retrieved October 28, 2021, from https://www.tripadvisor.com/TripAdvisorInsights/w3680.

What is Yelp's review solicitation penalty? | Support Center | Yelp. (2021). Retrieved 8 November 2021, from https://www.yelp-support.com/article/What-is-Yelp-s-review-solicitation-penalty?l=en_US

Why am I being told, "You're doing that too much…"?. (2021). Retrieved 7 December 2021, from https://www.reddithelp.com/hc/en-us/articles/204579879-Why-am-I-being-told-You-re-doing-that-too-much-

Wu, Y., Ngai, E., Wu, P. & Wu, C., 2020. Fake online reviews: Literature review, synthesis, and directions for future research. Decision Support Systems, 132, p.113280.

Zhang, J., Carpenter, K., Ko, M. (2013). Online astroturfing: a theoretical perspective Proceedings of the 19th Americas Conference on Information Systems, Vol. 4, Association for Information Systems, pp. 2559-2565

Zhang, K. (2020). How To Detect Fake Online Reviews Using Machine Learning [Blog]. Retrieved from https://scoredata.com/how-to-detect-fake-online-reviews-using-machine-learning-2/

Zhang, K., Cheung, C., & Lee, M. (2014). Examining the moderating effect of inconsistent reviews and its gender differences on consumers' online shopping decision. International Journal Of Information Management, 34(2), 89-98. doi: 10.1016/j.ijinfomgt.2013.12.001

Zhu, L., Li, H., Wang, F., He, W., & Tian, Z. (2020). How online reviews affect purchase intention: a new model based on the stimulus-organism-response (S-O-R) framework. Aslib Journal Of Information Management, 72(4), 463-488. doi: 10.1108/ajim-11-2019-0308

APPENDIX 1 QUESTIONS USED TO GUIDE THE INTERVIEW (ENGLISH TRANSLATION)

English translation of the interview questions

Intro
- General opinion about false reviews. False reviews = reviews that are not truthful
- What review platforms are in use?
- What social media platforms are in use?
- Prior experience regarding false reviews
- Opinion on why false reviews exist

Deterrence:
- Have false reviews been publicly brought up in the company media?
  - Internally?
- What factors deter false reviewers from trying?
  - Responsibility: the company vs. the platforms?
- How do the platforms deter false reviews?
- **Ideal vs. the current situation**

Prevention
- What has been done to prevent the possible damages from false reviews?
- Would the company reputation hold a momentary drop in rating?
- What are the platforms doing to prevent false reviews?
- Do the aforementioned factors deter the possible false reviewer?
- **Ideal vs. the current situation**

Detection
- To what degree can the reader detect the false reviews?
- How actively are the social media and review platforms used?
- What are the platforms doing to detect false reviews?
  - Is it working?
  - Differences between platforms?
- To what degree does the risk of getting caught deter the false reviewer
- **Ideal vs. the current situation**

Remedies
- What is the current plan of action for false reviews?
  - How would you act?
  - Recovery from the situation?
- Responding to a suspected false review
  - How?
  - What factors are focused on?
- How reliable is the report function present on the related websites/platforms?
- Does the rating returning to normal deter the false reviewer from trying?
- Should false reviewing be punishable
  - What type of punishments?
- **Ideal vs. the current situation**

RQ2: what factors deter from buying
- What factors prevent from buying positive reviews or negative reviews
- What would happen if you were to get caught from buying false reviews
  - Effects?
  - Experiences?
- How are the sellers of these services seen as?

APPENDIX 2 QUESTIONS USED TO GUIDE THE INTERVIEW (ORIGINAL FINNISH VERSION)

Haastattelukysymykset suomeksi
Intro
- Yleinen mielipide valearvosteluista. Valearvostelut = ei tosia arvosteluita.
- Mitkä arvostelualustat ovat käytössä
- Mitkä SoMe alustat käytössä
- Aiempi kokemus valearvosteluista
- Miksi valearvostelut syntyvät

Pelote:
- Onko valearvosteluita otettu julkisesti esille yrityksen somessa esim.
  - Yrityksen sisäisesti?
- Mitkä asiat pelottavat/estävät (eivät edes yritä) valearvostelijoita
  - Yrityksen oma vastuu vai some/arvostelualustat
- Miten mielestäsi some/arvostelualustat estävät valearvosteluita tapahtumasta
- **Ideaali vs. nykyinen tilanne**

Ennaltaehkäisy:
- Mitä on tehty valearvosteluista (mahdollisesti) syntyvän harmin estämiseksi?
- Kestääkö yrityksen maine valearvosteluista syntyvän arvosanan laskun?
- Mitä mielestäsi some/arvostelualustat tekevät valearvosteluiden estämiseksi (**estää julkaisemasta**)
  - Onko se mielestäsi tarpeeksi?
  - Eroja alustojen välillä
- Lannistavatko edellä mainitut asiat mahdollista valheellista arvostelijaa?
- **Ideaali vs. nykyinen tilanne**

Havaitseminen:
- Missä määrin lukija pystyy mielestäsi tunnistamaan valearvostelut
- Ovatko some/alustat aktiivisessa seurannassa?
  - Mikä saisi epäilemään valearvosteluita?
- Mitä some/arvostelualustat mielestäsi tekevät valearvosteluiden havaitsemiseksi
  - Toimiiko se?
  - Eroja alustojen välillä?
- Missä määrin kiinni jäämisen riski lannistaa valearvostelijoita?
- **Ideaali vs. nykyinen tilanne**

Korjauskeinot
- Millainen toimintasuunnitelma on valearvosteluiden varalta
  - Miten toimisitte?
  - Tilanteesta toipuminen
- Epäiltyyn valearvosteluun vastaaminen
  - Miten?
  - Mihin tekijöihin keskitytään
- Miten luotettavana toimena kommentin/arvostelun ilmianto alustoilla/somessa nähdään?
- Koetko että valearvosteluista palautuminen lannistaa valearvostelijoita yrittämästä (lopputulos +-0)
- **Ideaali vs. nykyinen tilanne**
- Pitäisikö valearvostelun olla rangaistavaa
  - Rangaistukset

RQ2: Mitkä asiat estävät valearvosteluiden käyttämisen
MItkä asiat estävät valearvosteluiden käyttämisen oman arvosanan korottamiseen tai kilpailijan arvosanan heikentämiseen?

- Mitä seuraamuksia kokisit, että valearvosteluista kiinni jäämisestä seuraisi
  - o Vaikutuksia?
  - o Kokemuksia?
- Miten valearvosteluiden/tykkäysten tarjoajat nähdään.