Consumers’ acceptance of digital referrals

Jonas Nikkonen

Master’s Thesis in International marketing
Supervisor: Malin Brännback
Faculty of Social Sciences, Business and Economics
Åbo Akademi University
Åbo 2022

Subject: International marketing	
Writer: Jonas Nikkonen	
Title: Consumers' acceptance of digital referrals	
Supervisor: Malin Brännback	

Abstract:

The purpose of the study was to research consumers' attitudes towards digital referrals by using the AEWOM model. The purpose has been researched by using four different research questions and three hypotheses. The research questions that were used in the study were "what are referrals in marketing?", "How and why do consumers rely on referrals?", "What factors influences acceptance of electronic word of mouth?" and "To what extent can the variance in acceptance of electronic word of mouth be explained by these factors? (Which of these factors is the most influential?)". These questions were answered in the literature review and in the empirical part of the study. The research questions supported the study and gave it a clear guideline to follow.

The literature review indicates that there are different kinds of traditional and digital referrals that can impact the consumers behavior and acceptance. Word-of-mouth (WOM) is the most common traditional referral, and plenty of companies have relied their marketing majorly on it. In general, consumers still use and rely on recommendations they receive in form of WOM, but with digitalisation a significant part of WOM has moved online, creating EWOM (Electronic word-of-mouth). Influencers, Recommendation systems and EWOM are the most common modern referrals from which consumers receive information about products and services. Consumers' attitudes towards the usage of referrals are quite evenly divided both for and against them. Many consumers feel that referrals in general help them save time and make decisions by narrowing the options, while others want to keep their autonomy regarding decision-making by avoiding any external manipulation which could lead to impulse purchases.

The hypotheses that were used in the study assumed that message credibility, source credibility and tie strength have a positive impact on consumers' acceptance on electronic word-of-mouth (AEWOM). To answer the hypotheses a quantitative study was implemented together with a multiple regression analysis. The independent variables that were used for the analysis were message credibility, source credibility and tie strength, while AEWOM was used as the dependent variable.

The findings of the thesis indicate that message credibility, source credibility and tie strength have a positive impact on AEWOM. The variable that impacted AEWOM the most is tie strength before message credibility and source credibility. However, all the variables were close to each other and therefore there cannot be claimed to be a clear difference between the variables. The findings were also compared to another study where same hypotheses were used, and same questions were asked in the questionnaire. The comparison between the two studies gave them both more credibility because the results were quite similar, showing that tie strength is affecting AEWOM more than the other variables.

Keywords: Digital referral, recommendation, Consumer, WOM, EWOM, AEWOM,	
Date: 20.4.2022	Number of pages: 62

Table of contents

1. Introduction and problem 1
1.2 Purpose 3
1.3 Research questions 3
1.4 Focus and delimitations 3
1.5 Methodology 4
2 Referrals in marketing 4
2.1 From WOM to EWOM 4
2.1.1 Influencers 7
2.1.2 Recommendation systems 7
2.1.3 Challenges 11
2.2 Data and AI & Cookies and SEO 14
2.2.1 Need for autonomy and GDPR 19
2.2.2 Amazon case 20
2.3 AEWOM 23
3. Research method and design 27
3.1 Quantitative research method 27
3.1.1 Hypotheses 29
3.1.2 Questionnaire 31
3.1.3 Respondents 33
3.2 Method 33
3.2.1 Multiple regression analysis 34
3.2.2 Reliability & validity 35
3.4 Substantial coefficients and variables 36
3.4.1 R^2 and adj. R^2 37
3.4.2 Beta coefficient 37
3.4.3 P-value 38
4 Analysis & Results 38
4.1 Descriptive statistics 39
4.2 Perceived task value 43
4.3 Hypotheses and research questions 46
4.4 Collinearity diagnostics 47
4.4 Analysis of Variance 48
4.5 Pearson correlation coefficient 49

5. Discussion 51
6. Conclusion 53
References 56

FIGURES

1. Information cocoon 13
2. SEO crawling 18
3. The first recommendations for Amazon in early 2000s 23
4. AEWOM model 26
5. AEWOM model including hypotheses 30
6. Message credibility 41
7. Source credibility 41
8. Tie strength 42
9. AEWOM 43

CHARTS

1. Descriptive data 39
2. Cronbach's alpha 44
3. Item statistics 45
4. Relationship between independent and dependent variables 46
5. Dependent variable explained by independent variables 46
6. Multicollinearity 48
7. Analysis of variance 49
8. Pearson correlation coefficient 51

Abbreviations

AEWOM	Acceptance of electronic word-of-mouth
AI	Artificial intelligence
CA	Cronbach's alpha
EWOM	Electronic word-of-mouth
MC	Message credibility
PPC	Pay-per-click
RS	Recommendation system
SEM	Search engine marketing
SEO	Search engine optimisation
SC	Source credibility
TS	Tie strength
WOM	Word-of-mouth

1. Introduction and problem

Consumers have always had the possibility to speak out and let others know if a service or product has been good or bad. By doing this, they have been able to help others make good decisions and support businesses that they think deserve it. Referrals have existed in different forms and especially with the assistance of new technology, new ways to recommend products for consumers have developed. Traditional referrals such as word-of-mouth are still valuable marketing sources for any company, even if a significant part of it has moved to social network. (Bughin, 2010) Artificial intelligence is the new technology that has made it possible for companies to recommend products for consumers by collecting users' information and preferences. These types of recommendations are created with recommendation systems, which have become a crucial part of many companies' business and will remain so, at least in some form.

The recommendation systems can be used with different platforms such as email, social media and online shopping web portals. The systems recommend relevant content that consumers could be interested in based on their behaviour and preferences. A recommendation system can collect information from websites, products, books, films and even songs, and with the help of algorithms, it will offer predictions and recommendations. (Behera et al., 2020, 1) The recommendations make consumers' buying decisions easier and more efficient. However, the recommendations can also impact consumers' autonomy and therefore be experienced as negative. (Quentin et al., 2017, 5) The recommendations are among the most common subcategories of AI, especially in B2C marketing and because of this their use evokes various emotions and opinions.

The usage of referrals in social networks have shared opinions both in favour and against it, as long it has existed and there are a few reasons for this. Most who experience referrals as positive have noticed different factors that help them in purchasing decisions, making their life easier. These factors are usually related to time-saving regarding product searching and decision-making. Many consumers also live a hectic everyday life which is why they value the usage of their time very highly, especially their spare time. The fact that there are more options available than ever before makes the search process and decision making even more difficult.

Those who are against the idea of companies using referrals in social networks are often worried about their autonomy in decision-making. In general, consumers want to make their decisions by themselves and without any external help. Modern referrals can impact consumers' decisions and lead to impulse purchases, which is something that many can find manipulative and unwanted.

Previous studies have indicated that many can see the value that recommendations can have regarding information search and comparison of different products. Most consumers still want to make the decisions by themselves without receiving any external assistance. (Vuolle, 2019) In a study by Niko Vuolle (2019), the results show that age and gender do not affect the need for autonomy among consumers. Other studies have shown a rising demand for customised customer experience where up to every third consumer thinks that brands are not focusing enough on customisation. (Komulainen, 2018, 301) Customisation can be done by targeting advertisement to a consumer according to how he has interacted previously, or by recommending different products that consumers would most likely be interested in. (Quentin et al., 2017, 15)

The way digital literacy and social network advertising are perceived has also partly to do with how consumers react to the referrals. Social identity is commonly considered to be one factor that impacts the way consumers accept advertising. Negative experiences, irrelevant content and sceptical attitudes towards the media or the message are as well factors that can be counted in the same group. Another phenomenon that has affected consumers' behaviour is fake news since it has significantly influenced their trust towards advertising on social networks. The 2016 U.S presidential election will be remembered for the fake news that was spread across social networks (Luna-Nevarez et al., 2015) and more recently this has also happened for news related to COVID-19. The acceptance of electronic word-of-mouth (AEWOM) is a model that can be used to see which factors affect consumers' acceptance of different referrals. Therefore, I have also used AEWOM in my thesis to help answering the research questions and to understand better what impact consumers' acceptance of information.

1.2 Purpose

The purpose of this study is to determine consumers' attitudes towards referrals by using the AEWOM model. As mentioned, the previous studies have shown different results where consumers are both for and against the idea of using referrals to support their buying processes. With this research, I also aim to gather as much knowledge of referrals as possible to be able to give advice to any company that would want to start using referrals as a support in their daily business.

Artificial intelligence (AI) and AI-based referrals have been popular topics for many years and several films and documentaries have been produced to help consumers understand how data are being collected and how their behaviour is being tracked. The increased knowledge among users can possibly lead to different results compared to any previous studies.

1.3 Research questions

The first and second research question will be answered in the literature review, while the third and fourth research question will be answered together with the hypotheses in chapter 4.3.

1. What are referrals in marketing?
2. How and why do consumers rely on referrals?
3. What factors influences acceptance of electronic word of mouth?
4. To what extent can the variance in acceptance of electronic word of mouth be explained by these factors? (Which of these factors is the most influential?)

1.4 Focus and delimitations

This study will focus on the theory of referrals and concentrate on consumers' acceptance and attitudes towards the usage of referrals in marketing. Referrals can nowadays be found in traditional forms such as word-of-mouth (WOM) as well as in more modern forms such as influencers and recommendation systems. To keep the theory in line with the purpose of the study, I will focus less on companies' perspectives and the technical factors of using the recommendation system itself and instead focus more on the reasons why consumers rely on referrals and what makes them accept the actual message or even forward it to their own network.

1.5 Methodology

The empirical part of the study will focus on collecting answers from consumers regarding their own attitudes towards the usage of referrals. The theory has been gathered from relevant books and articles and the purpose has been to use both newly published material and older material to support the theory in the best possible way. The method that will be used for the empirical part of the study is the quantitative research method because it allows collection of data on a large scale, which also makes the result more reliable. The questions used in the questionnaire have been used previously in another study with different respondents and in different contexts. The results of these two studies will be compared and discussed later in the thesis. The method will also be discussed more in chapter 3.

2 Referrals in marketing

In this chapter will be presented the most common referrals that are used in marketing, both traditional and modern solutions. The referrals that are mentioned in this chapter are not the only referrals available, but they can be considered interesting and relevant for the study.

2.1 From WOM to EWOM

Referrals have always had a great impact in customer buying behaviour and this impact has just grown with digitalisation. There are different kind of referrals that can be used to affect consumers' buying behaviour and to increase the possibility for sales. These referrals can be found in different forms, but arguably the most traditional and common referral is called word-of-mouth (WOM). It has been the main referral before digitalisation, and it has still managed to keep its role as one of the most important recommendation forms for companies to use. (Bughin et al., 2010)

When consumers decide to use WOM, it means that they have an opinion about a company's products or services that they want to share with others. The information can be either positive or negative, but it is still free advertising that originates from customer experiences. (Hayes, 2021) The reason why WOM is so effective is that consumers value others' opinions that are expressed directly to them and the consumer-to-consumer nature makes it very persuasive and credible. Even if the method is free, it can still bring more value than an expensive advertising campaign. (Bughin et al., 2010)

Social media and the other social platforms have made WOM even stronger than before and created a new sort of WOM, namely EWOM. Today, most of the word-of-mouth actions can be found online in electronic channels and this has increased its impact greatly. The EWOM-reviews are visible to everyone, and they can be found when searching for information about a company or its services. (Alboqami et al., 2015, 346)

Typical channels for eWOM are social media and blogs. These channels are used because they are free of charge and sharing information is easy. Facebook, Twitter and LinkedIn are examples of social network channels designed for sharing information. They all have their own user group, which also determines what type of content should be shared. Some consumers can go as far as creating their own websites or blogs just to praise or punish a brand. (Bughin et al., 2010) Social network channels and blogs allow users to chat with others and compare opinions.

EWOM and WOM are often used in the industry of hospitality and tourism where consumers tend to rely on previous experiences from other users before making their own decisions. These industries are selling services that are difficult to evaluate in advance, which means that reviews from strangers are considered important. In the article written by Cantallops and Salvi (2014), the authors explain how less-known hotels are doing everything to receive reviews because in so doing they differentiate from their competitors. The authors also indicate that negative comments do not have a significant impact on consumer behaviour. They explain that consumers mainly look at the overview of comments to see if they are positive or negative. A positive overview along with numerical ratings is a strong indicator for consumers. (Cantallops & Salvi, 2014, 47-48)

As mentioned previously, not all types of EWOM are considered equally crucial for consumers and their purchase decisions. Zhang et al. in their study "When does electronic word-of-mouth matter" (2010) claim that the consumption goals consumers associate with products can strongly influence reviews. Consumers often value positive reviews more highly and pay less attention to negative reviews, especially for products associated with preventive consumption goals. Negative reviews are good to some extent because they bring credibility to positive reviews. (Zhang, Craciun & Shin, 2010, 1340-1341) Ye et al. (2009), also mention in their study that positive online reviews can significantly increase the number of hotel rooms booked, while negative reviews have been proved to decrease sales. According to the authors, a 10% discrepancy in reviews can already improve the sales by 4.4%, and a 10% percent review variance can decrease the sales up to 2.8%. (Ye, Law & Gu, 2009, 181)

While there are several similar studies showing the impact that EWOM has in the most common industries, a different study shows the same impact also in the video game industry. The study results show that the impact of EWOM is different depending on the level of violence of the game. Those who play more high-violence games react more negative to review volumes than players of less violent games. However, there is no difference between players of single-player games and socially interactive games. The study shows how broadly researched EWOM is as a subject. (Zhang et al., 2019, 9)

Today, anyone can recommend a company to their own social network and earn some benefits from it. Many companies know how valuable referrals are regarding new customer acquisition, which is why new solutions have been developed. Companies can offer gift vouchers or cash to all customers who succeeds in attracting new customers through their recommendations. Arbatskaya & Konishi (2016) mention DIRECTV as an example of a company that offers a $100 credit to all customers who successfully obtain a friend to sign up for their services. Other companies where similar benefits can be offered are typically operating in the industry of banking, home alarm systems and housing. These referrals are often beneficial to all parties because the company gains new customers and both the current and new customer receive some benefits. (Arbatskaya & Konishi, 2016, 35-36) Other parties that are commonly relying on referrals are companies that are recruiting. Many companies are namely relying on their employees to introduce them to new candidates. (Mintz, 2005, 1-2)

2.1.1 Influencers

Influencers are a new type of referrer that have become popular due to the existence of social media and the problems that companies have with traditional WOM. An influencer's goal is to cause a reaction in someone else, as for example, a parent would like to impact the behaviour of a child, or a business owner would like to impact retail trends. Influencer marketing has its roots in WOM because it was originally viewed as a genuine endorsement and not an advertisement of the company itself. Marketers realised that recommendations were more credible when using influencers, and consumers were also more receptive to advertising in this format. (Brown, 2013) Influencers have been used as a marketing tool for several years, which is now possible due to right tools and platforms. Usually, they have also been a part of the target group as prospective buyers and that is why companies have not targeted them as a separate group. (Brown, 2008, 7)

Influencers are usually seen in the same channels as EWOM, but videos in particular are considered as the best format for social media influencers. The videos can in the best scenarios reach millions of viewers and if a beauty professional shows how to obtain the perfect bronzer, it will automatically affect the sales of the product. The best influencers are usually found by different agencies, so it can be difficult for brands to work directly with them. (Sammis et al., 2012, 12)

2.1.2 Recommendation systems

The number of digital data is increasing all the time, and this means that there is also a bigger demand for properly working recommendation systems (RS). RS can help a customer find and select the right products, which usually improves the decision making and the quality of the customer journey. For companies this usually means increase of sales. (Isinkaye et al., 2015, 262)

The modern referrals are supported by artificial intelligence and used to help companies gather information about their target group and to recommend personalised content to an individual customer. One of the most preferred solutions for this kind of referrals are the recommendation systems. RS is a software tool that recommends different items based on information learnt from consumers preferences. (Fesenmaier, 2006, 35-38) RS can help companies increase the average cart value as well as the customer engagement, and it can also be used in different channels such as email, social media, applications, televisions, and websites. (Behera et al., 2020, 3)

As mentioned above, the idea behind recommendation systems is to recommend items that online customers would be interested in purchasing. The user preferences can be collected from different sources as for example books, websites, songs and applications. With the help of an algorithm, RS can use the information to suggest predictions and recommendations to the users. (Núñez-Valdéz, 2012, 1186) Companies that decide to use RS usually want to offer their visitors as personalised e-shopping experience as possible, including both relevant suggestions of items to buy and several touchpoints during the shopping process. A failure with recommendation systems can lead to customer dissatisfaction and negative WOM. (Behera et al., 2020, 2)

A study from Infosys (n.d.) shows that almost one-third of the consumers that replied to the Infosys questionnaire wished they could experience more personalisation during shopping. 32% of the respondents had received product recommendations online and only 18% in-store. 20% claimed that they had never experienced any personalisation during previous purchases. (Infosys, n.d.)

Recommendation systems have been implemented successfully by several companies, for example Amazon, iTunes and Netflix. They all use recommendations to promote their products or services and to increase both sales and the time consumers spend with their services. Amazon uses the recommendation systems for product recommendations, iTunes for music recommendations and Netflix for film and series recommendations. This can help consumers save time and make their everyday choices easier. (Behera et al., 2020, 15)

Netflix is a company that has invested a lot of money to obtain a recommendation system that would meet its needs and expectations. Before the company's breakthrough, Netflix decided to offer one million dollars to a person or a team who would develop a better algorithm than what the company was using. It took more than three years before anyone managed to beat their algorithm in September 2009. The challenge included published data and closer to half a million ratings for 17.000 films. Netflix tested how the algorithms managed to predict ratings with the information from their datasets. A measurement method called RMSE was used to measure how well the algorithms performed. The algorithm that managed to win the price was a combination of different algorithms. (Leskovec et al., 2014, 336-337)

The internet is full of useful data that algorithms can use to optimise the information, but because the number of new data is increasing everyday it can also be a problem for recommendation systems. Depending on the data and the purpose for using RS, it can also affect the choice for which type of RS to use. Different types of filtering paradigms can be used to generate information, and the most common one's are content-based, collaborative system and hybrid approach. *Content-based recommendations* are recommendations made based on a person's past shopping behaviour or feedback. (Núñez-Valdéz, 2012, 1187) Different items are compared with items that the user has already evaluated, and in this way, it is also easier to recommend potential products to the user. However, content-based recommendations can only recommend similar products and therefore cannot adapt to changing preferences. Also, a relatively large amount of information and ratings are needed to recommend products based on the collected data.

The collaborative system uses data from previous purchases by similar consumers and makes predictions and suggestions based on this information. (Núñez-Valdéz, 2012, 1191) If a person has shared the same interests with another person in the past, the system assumes that this will be the situation in the future. The system needs a lot of information about the user's interests, behaviour, and activities in order to make accurate recommendations. Collaborative systems can be divided into two different types, memory-based and model-based. A memory-based system requires user ratings to predict the similarity between products and users, which is eventually used for recommendations. The model-based system relies on data mining and machine learning algorithms, which it needs to generate recommendations. The biggest advantage of a model-based system compared to a memory-based system is that it can adapt and handle different types of problems better. (Bansal, 2019, 30-36)

Hybrid approach is a combination of both content-based and collaborative system. (Núñez-Valdéz, 2012, 1193) Hybrid approach can reduce sparcity, scalability and cold-start problems by combining different sides from the two previously mentioned systems. Different algorithms are also made to generate with the system, which make it easier to provide accurate recommendations. Netflix is an example of a company that takes advantage of the hybrid approach as its recommendations are based on search comparisons and watching habits of both other users as well as one's own watching history and ratings. (Bansal, 2019, 40-42)

Some other recommendation systems that are worth mentioning are group recommender system, social network-based recommender system and context-aware recommender system. *Group recommender system* can filter information and help a group of consumers that consume items with each other. This can be needed in a situation where several consumers are planning a dinner or a vacation together. The typical approach for this system is to allow the users to engage in steering the recommendations.

Social network-based system includes data from the user and the social connection. It finds similar users by the help of tags, comments and co-authorships. Even if two persons share the same interest in films, they may still have different opinions on music, which is why category-specific groups need to be created and maintained. A recommendation system based on a context network is usually used for film recommendations as a complement to the user rating matrix. It can gather more relevant information about users and provide more accurate suggestions. (Bansal, 2019, 44-45)

A company should consider several aspects before implementing its own referral system. Cold start is something that is important to avoid, since it means that a company cannot recommend products due to a small amount of data. Algorithms can avoid such situations by recommending products based on their colour or price, but these types of recommendations are usually only used for seasonal sales or promotions.

Filtering rules can be integrated according to the consumer lifecycle. It can for example be very useful to offer social recommendations for a person that visits the company website for the first time. By doing this the person will receive recommendations based on his Facebook profile or another social network service. When implementing recommendation systems, companies tend to forget that RS can be used already at the welcome page and not only in the shopping basket when consumers are already close to a purchase. It can help companies increase the sales and obtain new customers, who would have otherwise made the purchase elsewhere. (Kembellec et al., 2014, 59-60)

2.1.3 Challenges

The usage of traditional and modern referrals includes plenty of challenges. As mentioned in previous chapters, most of the traditional WOM has moved online making it EWOM. According to previous research, approximately one-third of all WOM actions are active recommendations, which shows how important role it still has. (Alboqami et al., 2015, 345) However, the most significant problem with WOM is that companies cannot be sure if consumers have recommended their services to others, even if a person truly likes everything about the company. Also, obtaining referrals on a grand scale is very difficult and it is seldom something that a company can completely rely its marketing on. (Sammis et al., 2015, 27)

EWOM has solved many of the problems that has occurred with the usage of WOM, but there are also new issues that have emerged with the usage of EWOM. The communication is completely electronic, meaning there is no face-to-face communication, and this is something that consumers can find as less trustworthy. Moreover, not all consumers are looking for the information that EWOM can provide, which means that these consumers are also likely to ignore it. (Jalilvand et al., 2010, 44-45) Almost all reviews can be submitted anonymously, which means they are not as credible as WOM in general. According to Park & Lee (2008), the effect of negative EWOM has also a greater influence on consumers buying decisions than positive, and this means that even fake comments by anonyme users can harm a company's business and reputation significantly. (Park & Lee, 2008, 64-65)

Most companies are aware of the impact that EWOM can have on consumers buying decisions and this has also increased the chance for fake reviews. Sherry He et al. (2021) have studied the subject and according to the authors several companies are buying these reviews that are visible for example on Amazon. The issue that is usually being solved with fake reviews is the cold-start problem, that especially new products often have. The usage of fake reviews is not only doing harm for consumers, but also for other companies that are not involved in any cheating. The existence of fake reviews decreases the general credibility for all reviews, also for the ones that are real. Amazon has invested over $500 million and employed over 8.000 employees to decrease possibilities for any cheating and abuse of its platform (He et al., 2021, 42-43)

The usage of influencers can include a few challenges that everyone considering using influencers as a marketing tool should be aware of. Today it is common that companies collaborate with social media influencers to increase EWOM among consumers. Many factors suggest that it is worth using influencers as a tool to increase brand awareness, build emotional connections with consumers and maximise the reach of campaigns. However, the challenge regarding influencers considers mainly factors that can have a negative impact among consumers, such as bad quality content and excessive commercial orientation. High-quality content and a feeling of genuine interest in the sponsored product greatly affects the credibility of the advertisement. (Zhou et al., 2021, 122)

Even if modern recommendations are in general popular, there are still aspects than can be considered as negative, especially when combining with social media. In YouTube the algorithm can recommend videos that are classified as bad content for children, but still, it will continue recommending the content to keep the person watching. One of the main issues is also the fact that the creators of machine learning algorithms know what they have done to build a working algorithm, but they do not always know what the outcome will be when they are in use. This phenomenon is called “black box problem” and the lack of control can make it even evil in some circumstances. There are also other examples of how the RS can have a negative impact for the users. For example, YouTube video recommendations have spread recommendations about controversial content, Facebook's ad algorithm has been accused of discriminating against gender and race, and Twitter's algorithm has promoted extreme political rhetoric among its users, according to several sources. (Zoetekouw, 2019, 4-5)

Algorithm bias is something that is used when describing the narrow content that recommendation systems and the algorithms can lead to. Another well-known algorithm bias has to do with those who create them. It was claimed that several developers created algorithms for specific desired outcomes, such as areas of different cultural backgrounds that could be assessed as high mortgage risk areas based on race alone. In addition, job ads were only placed for men, so women did not even have a chance to see the job ad. (Stinson, 2020, 5)

Recommendations can vary depending on what RS is used and what kind of site is visited. In some cases, the recommendations can be too narrow and over-specialised making the customer only see products that he would not even be that interested of buying. (Stinson, 2020, 5) The algorithm bias is usually considered to be unfair since it only favours a certain kind of product categories or content. Even if recommendations are not designed to be fair, many customers would still prefer balanced mix between the recommended products. (Zoetekouw, 2019, 4-5)

Homogenisation can also be a problem for consumers visiting sites with recommendation algorithms. Homogenisation can lead to over-recommendation for the most popular products, and this happens when algorithms are designed to first show the products that others have liked. This means that the entire dataset is not showed equally, and the supply becomes very narrow. (Stinson, 2020, 3-4) The term that is often used for this kind of situations is *information cocoon.* Recommendation systems will recommend products that are already in line with the users' preferences, which might sound like a positive factor in the short run, but in long term it might feel like consumers are being trapped in their old consuming habits. (Chen et al., 2021, 2-4)

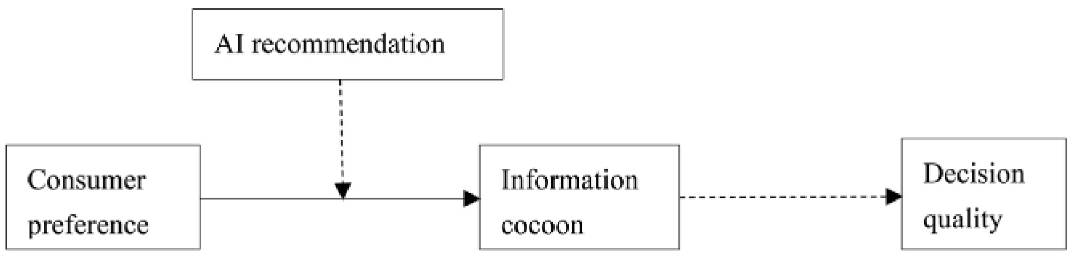


Figure 1: Information cocoon (Chen et al., 2021)
The figure shows how AI technology affects the recommendations for consumers. The usage of information cocoons can be helpful for many consumers, but it also shares opinions because of its manipulative reputation.

Several concerns have to do with data leaks to a third party or even with data gathering without permission of the user. It can also be seen as a threat if a third party receive information about given recommendations, since that can also be used as a tool for manipulation, without the user even being aware of the action. (Milano et al., 2020, 962) To avoid this to happen, it could be possible to have privacy controls where the user would decide if their own data can be shared and with whom. However, this choice is something that many consider unnecessary and as too big of a burden for consumers to take responsibility of. (Paraschakis, 2017, 215)

2.2 Data and AI & Cookies and SEO

Almost everything that consumers do can be measured, saved and analysed. When we visit different websites or communicate with each other on the internet, we also leave a digital lead that can be used in different ways against us. Typically, the data is used to make money and increase business in general, as well as to predict consumer behaviour. (Aaltonen, 2019, 161) Data can be in different forms, such as numbers, texts, images or even videos. It becomes information when there is a purpose that can be linked to it. Therefore, pure data is not always meaningful or valuable. The quality of data can vary. Therefore, it is advisable to use systems consistently and decide in time what kind of data is wanted. With this information it is also easier to develop the systems in the right direction. (Puolitaival, 2019, 72)

Many department stores have long offered rewards cards to their customers and have been able to use this data to improve their product development and product placement. However, these department stores have typically missed the social connection with customers on social media. (Rubanovitsch, 2018, 114-118) The most common sources of data collection are sensors placed in various machines, applications, social media, mobile phones and search engines. The data provided by sensors usually relates to energy consumption, production consumption or consumer purchasing behaviour. (Aaltonen, 2019, 94)

The data that comes from applications is usually taken from Customer Relationship Management (CRM), Enterprise Recourse Management (ERM) or from a corporate website. By using social media, it is possible to collect a considerable amount of data from all different users. Various social media tools help to analyse the data so that it can eventually be used in marketing and decision-making. Mobile phones and their applications can also provide a considerable amount of data, such as time and location, and most importantly, the applications can identify what a person prefers or is interested in. This information is something that most marketers and analysts need.

Search engine data can be received from the search engines such as Google Insight. Search engines have also created their own subcategory of marketing, namely search engine marketing (SEM). SEM is used for following what words consumers tend to use when searching for information and then utilise the information to end as high as possible in the search results. (Aaltonen, 2019, 95) Data is usually mentioned in the same context as information, which is a result of processing data. Data itself is usually not valuable, but when it is processed and can be used as information, it becomes useful. Data can also be divided into different types or categories. *Qualitative data* is non-numerical and can be as simple as a colour. *Quantitative data*, on the other hand, has to do with numbers, such as the number of animals or a quantity. *Binary* is a term that is not used as often, but it can stand for simple data such as "true/false", "yes/no" or "on/off". *Ordinal* stands for data that can be found in a ranked order, while *nominal* stands for the opposite. (Sanders, 2016, 2-4)

Artificial Intelligence has become one of the most interesting and discussed subjects in the 2010s, even if it was originally invented already in the 1950s. The development has taken massive steps forward from the early days, which is why it has spread to the consciousness of ordinary people. (Dick, 2019)

Multiple different definitions can be used to describe Artificial Intelligence, but one that is widely used is the definition from the artificial intelligence reporter Dave Gershgorn (2017). He has described AI as a software or a computer program with a learning mechanism. John McCarthy who is said to be the father of AI has described the process as “making a machine behave in ways that would be called intelligent if a human were so behaving” (Press, 2020). Even if most researchers have their own opinion of what AI really is, there is still one aspect that most of them can agree on. Namely, most researchers believe that how the problem is addressed is almost as important as whether it is solved. (Kaplan, 2016, 2-5)

AI is being used now more than ever, and businesses are no exception. Some of the recent studies show that most consumers do not know exactly what artificial intelligence does or how it affects them, even though a majority have interacted with the technology through chatbots, shopping recommendations and search systems. The reason for the popularity of artificial intelligence is because it has advanced so much in recent years, making it a serious option for every marketeer to consider. (Komulainen, 2018, 298-299)

The major investments that companies are placing in AI have led to more discussion around it and many are slowly starting to understand what it is all about. Google, Amazon, Facebook and Apple are examples of companies that have started to concentrate majorly on its consumer-oriented side. (Botha, 2019) Many discussions have suggested that robots will replace humans in several jobs in the future and therefore, it should not be a surprise that the technology will also be used to affect our buying processes and decisions that we make every day. (Sterne, 2017, 20-24) Artificial intelligence can be used in different ways to improve customer satisfaction and the whole experience of the company. Today, it is easy for companies to customise the whole buying process for each customer with data that can be collected by help of AI. (Komulainen, 2018, 300)

The definition of **cookies** can sometimes vary, but it is in general a small text file that is stored by a website that consumers visit with their electronic devices. The commonly seen “remember me” view that appears on our screens when we visit different websites, will place a cookie on the device until the cookie expires and that means that the device will automatically also remember the log-in details during this period. (Kaushik & Prakash, 2018, 458-459)

The main idea with cookies has been quite the same since the beginning of 1994, which is to provide better user experience and add more functionality for consumers when visiting different websites. However, there has also been clear development steps towards collecting data from the users, which users usually allow since they do not always understand the negative aspects that the choice can lead to. If a user ignores the safety instructions, this usually leads to decisions without being aware of the true reasons for their actions. (Kulyk, 2018, 2-4)

Even if cookies are used to collect user data, this can be considered a safe tool for the end user. The information collected can only be read by the server that is visited and it is only one of many tools that webmasters can use to check users' website activity. Cookies can be divided into two different groups, namely session cookies and persistent cookies. (Mitchell, 2012, 4)
Session cookies are helping websites remember the user information as for example what items have been added to the shopping cart and what other activities the user has done on the pages and in what order. *Persistent cookies* save the information from the session and stores is long after the session has ended. The information that is collected with cookies can be for example login details and user preferences. Collecting this kind of private information means that the cookies are usually encrypted by the web server before storing it on the users’ computer. (Mitchell, 2012, 5-6)

By using cookies, companies can target their marketing more efficiently and personalise it more effectively than before. In the study by Pelau et al (2020), can be seen that consumers are in general willing to share their personal information by accepting cookies, in change for discount or other benefits. The authors explain that consumers tend to believe that they have lower tendency to be manipulated by targeted marketing than others. (Pelau et al., 2020, 830)

Even if using cookies includes a lot of benefits for the companies, they have also some down sides. Cookie profiling can be seen as one of the down sides, since it can provide targeted advertisement cross platforms, systems, and services. This means that a person interested in buying a new car may receive targeted advertisements on various websites that are in no way related to cars, and the person may also be contacted offline through phone calls. This can easily go too far and result in a bad outcome for both the user and the company using cookies and advertising. (Kaushik & Prakash, 2018, 459-460)

Data protection is usually the main issue that concerns consumers when it comes to the use of cookies. In particular, the fact that third parties may in some way be able to use the information collected using cookies. This information could then be used to create an informative profile with all user preferences, which would not be possible without cookies. The reason why information can end up with a third-party provider is because online advertisers strive to create the best possible buying experiences, and for this they usually need third party tracking cookies. (Mitchell, 2012, 3-6)

Search engine optimisation is a free tool that companies use to improve their visibility on the internet. Without SEO, content marketing loses its meaning because without a tool to drive traffic to the site, there are no new consumers finding the site. The way companies handle the SEO process also affects their ranking in search results. Google shows the results in order of importance and usually the first links are the ones that are opened. The results that end up on the second page have already lost 75% of all consumers who searched for the specific information. (Komulainen, 2018, 150) It does not matter how good a company's service or product is if no one knows about it, and as the old saying goes, "If a tree falls in a forest and no one is around, does it make a sound?" (Bradley, 2015, 1-3)

SEO is in general a subcategory of SEM (Search engine marketing) and the common goal for all companies using SEO is to be the first one in search results. The most important factor that determines how a website is ranked is the search engine spiders that collect information from the website and rank the website based on the information they have. Factors that affect this ranking are loading speed, ease of navigation, placement of links that point to relevant information, and the quality of the website's content. Even if a website is visually well done, it may still have some technical flaws that the spiders do not recognise, which means that the site will not appear high up in the search results. (Carter et al., 2007, 145-146)

To understand how search engines are used, it can be profitable to understand what happens under the surface. Search engines must discover all the data available on the internet to be able to offer best possible results that match with the search query. Typically, search engines start by finding information from high-quality websites that have a good reputation, and then use the links on those websites to find other websites. Automated robots, also called spiders, can reach billions of documents using these steps. (Enge, 2012, 34-35) The search engine analyses the information it receives and may deliberately skip some pages that do not look enticing enough.

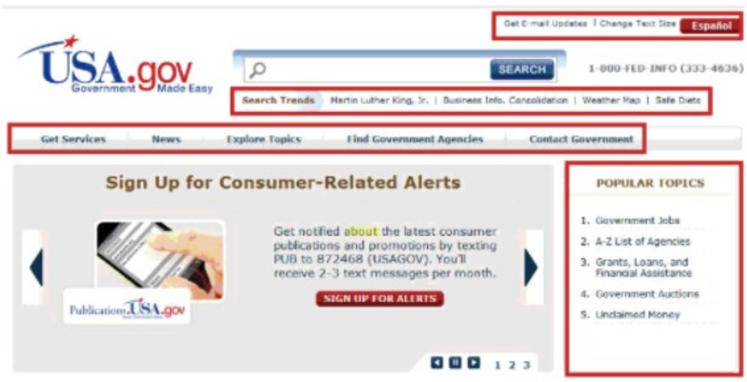


Figure 2: SEO crawling (Enge, 2012)

In the figure above can be seen how the search engine loads different pages and analyses them. This action can be repeated many times until every piece of important information has been found.

The codes that are collected from different sites are saved on hard drives where they can later be taken from if needed. All the collected data must be saved in a data centre constructed by search engines. An important aspect to consider when creating search engines is the starting point of the search or crawl. In general, it is advisable to start with a trusted website, as this makes it easier to compare how it reacts to less trustworthy sites. (Enge, 2012, 35-36)

Besides search spiders or search bots, there are also another category of SEO which is Pay-per-click (PPC). The usage of PPC begun in 2008 when ads started to appear in search results. The ads are shown on the top of the results, and they only cost when someone clicks on the links, which explains why they are called pay-per-click. The ads look like the organic results, and that is why consumer do not always realise how they differ from each other. The biggest difference between organic results and PPC is the text that is displayed next to the link. The text next to PPC links varies depending on the search engine used, but there is usually a small text for "ads" or "sponsored results". (Kent, 2012, 25-28)

2.2.1 Need for autonomy and GDPR

Consumers have never received as much information about different options as today, regardless the product or service. The referrals that companies tend to use can in best scenarios be helpful, but they can also have a negative impact on consumers' autonomy. Consumers value their autonomy in general quite high, which of course, raises questions about the use of recommendations. Consumers usually have a need for autonomy in decision making and this subject has also been researched from different perspectives such as psychology, philosophy, and consumer research. Autonomy is a way to experience free will and self-determination and even if a decision would have its routes from elsewhere, consumers still want to believe that it was their idea, and it is also common for them to find good arguments for unnecessary purchases. (Quentin et al., 2017, 27)

André Quentin et al (2017) mentions a study where the results indicate that consumers usually want to reinforce their self-will when they feel that their autonomy is threatened. The results indicate that when consumers understood that various algorithms made their choices predictable based on their past preferences, they began to choose options that they preferred less to leave their autonomy untouched. (Quentin et al., 2017, 430-433).

Autonomy and free will can lead to odd choices, but why is it so important for us? Autonomy is for example said to help consumers develop a moral responsibility. When continuity is experienced between our beliefs, actions, and their outcomes, there can also be experienced pride and closure. When the beliefs and thoughts are inconsistent with the actions consumers usually feel guilt and regret. When consumers experience making their own choices, this can also show up in the form of positive outcomes. For example, ordering healthier and less tasty food in a restaurant to show self-control can also send positive signals to the person and strengthen their willpower and autonomy. (Quentin, 2017, 32-33)

The discussions around customer autonomy have led to different ideas where consumers could for example modify the algorithms by themselves on the websites that they visit, or by using different technologies making it easier to crowdsource or co-create. (Quentin et al., 2017, 436) Lena Bjørlo et al (2021) have plenty of suggestions on how to improve the consumer autonomy and one of them would be the complementarity of AI technology that could be used as a value bringing tool leaving consumers still enough room for decision making. (Lena Bjørlo et al., 2021, 11-12)

GDPR or General Data Protection Regulation is a regulation that specifically affects businesses and organisations that collect data from consumers in the EU. The GDPR has been in force since May 2018 and the high fines have prompted companies and organisations, at least in part, to adapt to the regulation. The aim of the regulation is to make consumers safer, especially at a time when more and more consumers are sharing personal data with various online services. Personal data is generally any information that can be used to identify a person. For example: Name, email and addresses, but also gender, ethnicity and place of residence can be listed as personal data. (GDPR, 2021)

Even if the collection or storing of data is very restricted, it is still not completely forbidden. It can be justified when the permission has been given separately for example by opting into the marketing list. One other accepted reason can be when preparing a contract where the data subject is a part of. Sometimes it may even be sufficient as a ground if a person simply has a legitimate interest in processing an individual's personal data. The GDPR has made consumers more aware of their data protection rights, which are also divided into different categories. They have the right to be informed, the right to access, the right to object and the right to restrict processing. (GDPR, 2021)

2.2.2 Amazon case

Amazon is an excellent example of how companies can use recommendations in their business strategy. The company was founded in 1994 by current CEO Jeff Bezos who worked as a programmer at various trading firms at Wallstreet before deciding to start his own company.

He wanted to start his own business in retailing and decided to choose books as the focus area. Together with his team of ten, he developed a software and created a website, which was launched in 1995. The company began selling books to online customers, who only had to enter their email address, credit card and password to make the purchase. (Thomas, 2021, 301-303)

Recommendation algorithms can be used in a variety of ways. However, the most common way is to collect data about customer interests and then use this information to recommend products to potential customers. Many recommendation systems mainly use data of the bought products and given ratings to be able to recommend new products, but demographic data and subject interests can also be used for recommendations. Amazon started to personalise their online store for every customer making it look different depending on the person who visited the site. Those who were interested in software engineering received recommendations related to programming, while for example, women who had recently become mothers received recommendations related to children’s toys. (Linden et al., 2010, 76-77)

When Amazon began to attract significant number of customers, it also meant that the number of data increased with major volumes. The increase in data can be important for a company, but it can also be a burden because it makes it more complicated to review. For companies like Amazon, the most suitable recommendation algorithms can scale big amounts of data, react to possible changes in customer data and do not require many purchases from a customer to obtain reliable data. (Linden et al., 2010, 78-79)

Amazon has invested time and effort to enable the recommender systems that it uses today. Amazon’s algorithm uses less data space than most other algorithms, which is one of its biggest advantages it offers. Most of the company’s products reach end-users because of recommendations and this has been the situation from the beginning. Other famous companies known for the successful use of recommendation systems are for example Youtube, Spotify and Netflix. 30% of all visitors that Amazon receives are from recommendations and 80% of films/series watched on Netflix are also due to recommendations. This shows the significant role recommendations can have when tailoring a special service for customers. (McDonald, 2021, p. 10) Recommendations can also be compared with a store visit where all the shelves are in motion and only relevant products are displayed for each customer, helping the customer to find relevant products and save time, which usually means an increase in sales for the company. (Smith & Linden, 2017, 12)

The role of recommendation systems cannot be overemphasised, especially in product catalogues, where the main task is to ensure that the right content is displayed to the right consumers. Recommendations generally become so accurate that further approvements need to be carefully made to avoid inadvertently breaching the GDPR. Both Amazon and Google have claimed that they followed the regulations and therefore did nothing wrong. This kind of speculation shows how good a job Amazon has done with its recommendation system to even engage in this kind of discussion (McDonald, 2021, 15)

In addition to accuracy, timing is usually critical with companies and their various product catalogues. Amazon is constantly changing its catalogues, and this is also typical for companies that operate in the fashion or technology industry where changes and innovations can happen very quickly. Technology can be difficult because new innovations usually have a cold start before consumers read about the products in the media. Amazon has chosen not to limit its recommendations too much based on previous purchases. Companies that only sell one product category will naturally recommend similar products in the same category, but it is beneficial for a company like Amazon, to also recommend something to customers that they have not yet considered and something that they could benefit more in the long-term. (Smith & Linden, 2017, 13)

Amazon uses item-based collaborative filtering, which is known to be simple and easy to use. Recommended products are displayed based on previous interests. When the items that the person has already seen or bought are filtered out, only the products that the person might find interesting remain. Most of the work behind this algorithm is done offline, but when the information is processed offline, it means that the filtering can be seen online in real-time. The recommendations are high quality, which means that consumers are usually not offered as many irrelevant recommendations. The high-quality recommendations can be offered with a higher probability than most other systems because the algorithms are able to scale hundreds of millions of users and items without having to sample them individually, which can weaken the quality in some situations. (Smith & Linden, 2017, 15-16)

Before Amazon decided to use collaborative filtering, the world mainly focused on user-based recommendations, where users were compared to each other. The idea of focusing on similarities between products rather than other users proved to be the right one and something other companies had not yet done. Item-based filtering offered many advantages, some of which have already been mentioned. Amazon had the choice to compare visitors in real time and find the best matches, or to create an offline index where they could compare customers with each other. (Hardesty, 2019)

The problem with a significant number of different catalogues is that the purchase history changes for each customer, which also means that the index has to be updated daily. The method still requires less work than identifying customers who visit the website, and so it was with the technology from early 2000s. (Hardesty, 2019) Back then, consumers also had the option to click on the “Your recommendations” link to see the basis on which recommendations were made, filter their own view by choosing between products or subject areas and rate the products they bought. (Hardesty, 2019)

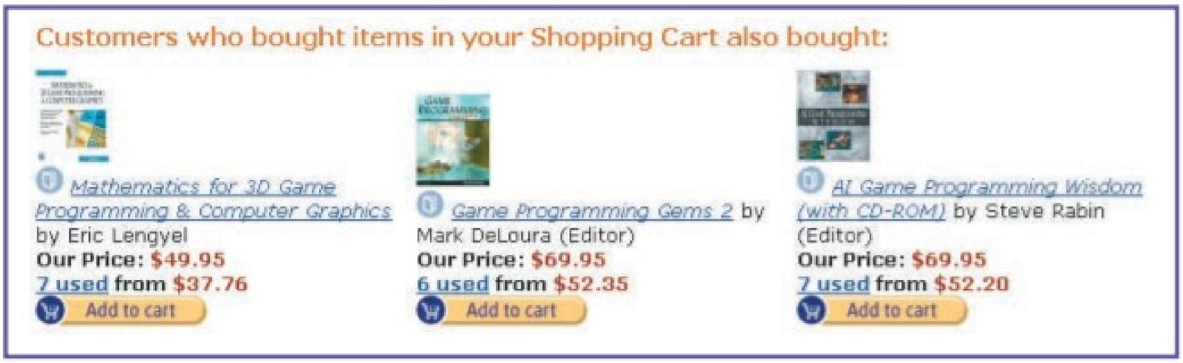


Figure 3: The first recommendations for Amazon in early 2000s (Linden et al., 2010)

The figure shows the first recommendations that Amazon started with. They used the items in the shopping cart to recommend similar products to the customer. It can be compared with impulse purchases in a grocery store before checkout, but these recommendations were tailored for each customer. (Linden et al., 2010, 78)

2.3 AEWOM

AEWOM (Acceptance of EWOM) is the theorical model that I have used in my study. AEWOM means the degree to which customers agree and accept online recommendations and comments. AEWOM has been identified as a key component to improve customer trust and build a strong brand. Customers can communicate with each other through EWOM and obtain more information about products and services, which enhances the process from perceived value of the product or service to consumer trust. (Ruan et al., 2020)

The value perceived by consumers evolves according to their self-awareness in the consumption experience, and EWOM involves the process of information intake in advance of accommodation. Once customers establish through actual experience that their perceptions match the positive content of online reviews, trust can be established. In addition, potential customers are more likely to search information from experienced customers before consuming because EWOM can reduce the risk and uncertainty of unknown consumption. the impact of AEWOM on customer trust is stronger when the customer has had a customer service experience and value development occurs. The acceptance of positive AEWOM also leads to a significant trust among consumers. In other words, AEWOM can act as a catalyst to accelerate the process from perceived customer value to trust. (Ruan et al., 2020)

The role of consumers in attracting new customers has increased over time, and this is also why companies focus on receiving referrals from their current customers. Sabita Mahapatra and Abhishek Mishra have written an article about "Acceptance and forwarding of electronic word of mouth" (2017, 594-595) and in their article they have among other things pointed out the most relevant information that should be known about AEWOM. Source credibility is one subject that is usually discussed when talking about AEWOM and that is also understandable when the comments are mainly submitted by strangers. However, because there is usually no financial profit as a motive behind the comments, they are seldomly questioned for trustworthiness. (Mahapatra & Mishra, 2017, 597)

A consumer is likely to accept the referral when he finds it credible enough in form of good content and perceived usefulness. When these are achieved it is likely that the message is accepted as well as forwarded. (Mahapatra & Mishra, 2017, 597) Other studies have shown that credibility is strongly associated with credible and valuable information and leads more likely to acceptance and forwarding of the message. (Cheung et al., 2009, 14-15)

Mahapatra and Mishra (2009) focused their research on source credibility, message credibility and tie strength. Their findings show that source credibility affects information acceptance and that strong ties usually lead to greater trust and dissemination of ideas and information, which in turn increases EWOM acceptance. The less significant relationship between credibility and electronic word-of-mouth indicates that consumers generally do not pay as much attention to the credibility of the message as they do to the credibility of the source and tie strength to justify the acceptance of online reviews. (Mahapatra & Mishra 2017, 602)

EWOM cannot be avoided completely by ignoring it. Online reviews are available everywhere which has made it easier for consumers to make decisions regarding purchase of products or services. Consumers tend to rely on others' opinions about a product or service, especially when they want to hear an impartial opinion. Many industries are known for having plenty of channels where consumers can submit reviews and have discussions with each other around these topics. Companies that operate in travel and tourism industries are especially known for providing services that are difficult to evaluate in advance, therefore, to facilitate the decision-making consumers tend to rely on online reviews submitted by others. Online reviews and discussions usually occur between consumers with weak social ties, which can affect the reliability of the message itself. (Alrasheedi et al., 2021, 2124)

According to Wang & Wei (2006) there are several aspects that lead to consumers acceptance of EWOM. Consumers need to collect information to gain knowledge to be able to form an opinion about the received information. The sources have a significant impact on the attitudes and opinions and therefore they are also a crucial part of the acceptance in general. Factors that can lead to low recommendation acceptance are for example situations where the consumers need for information and the EWOM product information result in low diagnostic assessment. (Wang & Wei, 2006, 782)

As mentioned previously, online reviews are typically submitted by strangers, which means that the tie strength is usually weak, therefore the information source can be quite crucial for accepting and believing the message. A message is commonly experienced as credible when the receiver finds it believable and factual. Style, attractiveness, quality, and source are factors that typically affect the acceptance and the attitudes towards a message in general. Acceptance of online reviews also tend to be likelier when both the receiver and reviewer have same type of background, interests, or life situation. Messages that are accepted as credible by consumers are more likely used for decision-making, while it is unlikely that non-credible messages are accepted or used for this kind of purpose. (Alrasheedi et al., 2021, 2128)

The factors that impact how consumers accept online recommendations can vary a lot depending on the cultural background, which means that the acceptance cannot be generalised to a few specific factors. According to Fan et al. (2017), the differences between cultures can be quite significant between the consumers with different cultural backgrounds. The authors explain that Korean consumers are more likely to accept online recommendations than consumers in USA and the impact on purchase decisions is also greater for Korean consumers. However, their study also shows that the difference between Chinese consumers is not significant compared to consumers in USA. The only difference that the authors could claim based on the study, is that Chinese consumers have a significant trust for others in their social network, but the situation is not the same regarding online recommendations submitted by strangers. (Fan et al., 2017, 150-151)

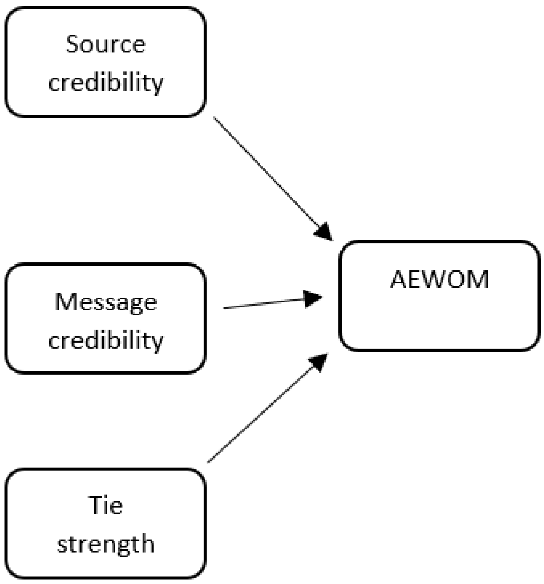


Figure 4: AEWOM model

The AEWOM model is used for measuring the acceptance of electronic word-of-mouth and the impact that source credibility, message credibility and tie strength has on it. The model can also be modified for example by adding forwarding of the message as one variable.

A similar model that can be used for the same purpose is called EWOM Information Acceptation Model (IACM). Javier Torres et al (2017) have written about the model in their article “Impact of gender on the acceptance of electronic word-of-mouth (EWOM) information in Spain". The variables that the authors have used for studying the acceptance of information are quality of information, credibility of information and their impact on usefulness of information, which leads to acceptance of information. (Torres et al., 2017, 2)

Technology Acceptance Model (TAM) is another model that has been used for a long time and it is proven to be one of the best models for other purposes, such as to describe an individual's acceptance of information systems. Besides Techology acceptance model, there are also a few other models that have been successfully used to understand human behavioural. Maybe the most famous ones are theory of planned behaviour as well as the theory of trying, that have been widely used to explain behaviour. (Bagozzi, 1990, 131)

3. Research method and design

This chapter will go through the research methods that are used in the study, introduce the hypotheses, and questionnaire as well as the target group that was chosen. A research method is generally needed to be able to collect data for the study. Before choosing the research method, it is recommended to read about the different options available before making the final decision. (Bryman & Bell, 2011, 41)

3.1 Quantitative research method

This thesis will include a quantitative research method in the form of a questionnaire because it supports the research questions in the best possible way. A questionnaire is suitable for this specific thesis since the research questions require it and personally, I also want to learn more about quantitative research methods. Questionnaire research is used for collecting data by a questionnaire or by several interviews at a single point in time. Otherwise, it would be more difficult to collect quantitative data connected with at least a few variables. (Bryman & Bell, 2011, 54)

According to Bryman and Bell (2011), the quantitative research method has in general been the most popular method in business research, even if the qualitative research method is also increasingly used. The quantitative method has been described as a distinctive research strategy and its main preoccupations are features such as measurement, causality, replication and generalisation. (Bryman & Bell, 2011, 150-151) Quantitative techniques makes it easier to study larger groups and draw generalisations from different samples. Quantitative methods offer multiple paradigms and approaches and without them, research would be more limited. The research technique itself can also be perceived as confusing due to different steps and paradigms involved in the method. Using quantitative research methods has been compared to learning a new language, which can be difficult at first but becomes easier over time. (Swanson & Holton, 2005, 30-32)

The quantitative research method is usually used after some theory has been collected and hypotheses have been made, which are later measured quantitively and analysed according to the research procedures. Data is collected mainly for two different purposes: to better understand a specific phenomenon under study and to draw conclusions from a wider group than the one under study. The research method can be considered strong when it comes to finding detailed information about a topic. However, like all methods, the quantitative method has its downsides, and the success of the research depends mainly on how well the method fits the purpose of the study. (Swanson & Holton, 2005, 30-32)

According to Bryman & Bell (2011), there are 11 main steps that anyone considering using the quantitative research method should be aware of, even though most studies do not require each step to achieve the desired outcome. **Step one** involves gathering relevant theories that can be used to better understand the results and gather knowledge about the topic. **Step two** involves hypotheses to be tested that are likely to be found in experimental research. **Step three** involves selecting the research design and **step four** involves developing measures of concepts the researcher is interested in, also called operationalisation. **Steps five** and **six** are about selecting research sites and subjects/respondents. **Step seven** is called administration of research instruments, which means pre-testing with different subjects or manipulating independent variables.

Step eight is about collecting data. In quantitative research, this usually means preparing the information so that it can be quantified. **Step nine** is about analysing the data. Here the data are tested with different variables and different ways of presenting the results of the analysis are developed. **Steps ten** and **eleven** are about developing and recording the results/conclusions. The hypotheses need to be compared with the results, and everything needs to be written down clearly. (Bryman & Bell, 2009 p. 150–153)

3.1.1 Hypotheses

A hypothesis explains possible outcomes of an experiment. The possibility of having no significant results is always there and therefore a null hypothesis (H0) can be needed, stating that nothing notable happened. In addition, an alternate hypothesis (H1) is needed to state that something significant did happen, which usually is the goal. When the data has been analysed and the results are clear, it will be easier to know which of the hypotheses should be accepted and which one rejected. (Knapp, 2018, 12-13)

I have used to some extent the same hypotheses as Sabita Mahapatra and Abhishek Mishra in their study "Acceptance and forwarding of electronic word of mouth" (2017). They have measured Message credibility with an eight-item instrument made famous by Beltramani (1982), source credibility with an eight-item instrument by Netemeyer (1999) and tie strength with a four-item instrument by De Bruyn (2008). Finally, AEWOM was measured with a four-item instrument made famous by Wu and Shaffer (1987).

The following hypotheses are used in the study:

H1. Source credibility is positively associated with AEWOM.
H2. Message credibility is positively associated with AEWOM.
H3. Tie strength is positively associated with AEWOM

H1. Source credibility is positively associated with AEWOM.

Source credibility is one of the factors that are considered to impact consumer acceptance towards referrals. Findings from previous studies indicate that consumers are likely to accept information from others who they believe to be homophilous against themselves. Willemsen et al (2012) studied how consumers would rate the credibility of different posts by posting product reviews from three different sources on a website. The reviews were placed by someone who claimed to have some knowledge about the product (layperson), someone who claimed to have work-related knowledge about the product (semi-pro), and someone who was a rating expert and had been evaluated by other users in the past.

The results of the study showed that self-proclaimed expert received high points in source expertise and the layperson high points in trustworthiness. However, the rated expert was the only one who scored the highest in both categories. (Willemsen et al., 2012, 508-510)

H2. Message credibility is positively associated with AEWOM.

Message credibility is referred as a review or recommendation that is considered believable and trustworthy. Usually, a message is considered trustworthy when the receiver finds it believable and factual. (Sweeney et al., 2012, 239)

H3. Tie strength is positively associated with AEWOM

Communication between peers on social media is dependent on tie strength, which is usually found on different forums and groups with close-kit communities. Tie strength is described as the relationship between the members of these groups, and it can range from weak to strong depending on several factors, such as emotional intensity, time spent together, and intimacy shared. (Granovetter, 1973, 1360-1361) Strong ties is commonly resulting in higher trust and effective transfer of ideas and information. (Hansen, 1999, 82) The closeness of peers is also increasing the acceptance of EWOM. (Mazzarol et al., 2007, 1486)

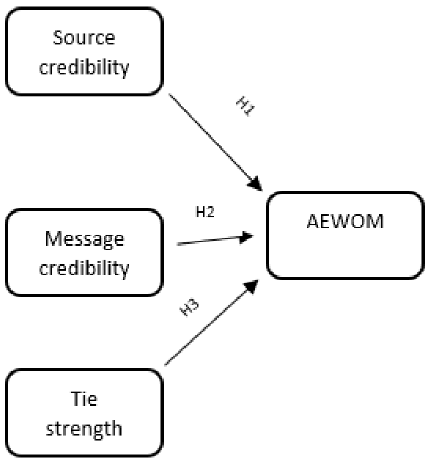


Figure 5: AEWOM model including hypotheses

The figure illustrates the hypotheses used in the study. The hypotheses assume that all the independent variables will have a positive impact on the acceptance of electronic word-of-mouth. The research questions used in the study also aim to answer questions regarding, which of the independent variables has the most influence on the dependent variable and to what extent can the variance of AEWOM be explained by the independent variables.

3.1.2 Questionnaire

Survey research is implemented with a questionnaire that is used for collecting data in connection with two or more variables. Internal validity is usually considered weak in this kind of research because casual direction can be difficult to establish from the results. However, external validity is usually strong when the data has been randomly selected from a specific sample. (Bryman & Bell, 2011, 54-55) The size of the sample is commonly argued because there is no specific number of responses that should be followed as an instruction. The absolute size is generally considered to be more important than the relative size, meaning that for example 1.000 individuals in UK will have as much validity as 1.000 individuals in USA, even if the total population is different size. A large sample usually leads to fewer sampling errors and that is why the number of sampling errors should be important to consider when making the decision of sample size. (Bryman & Bell, 2011, 187-188)

The sample size or quantity is commonly not as crucial as the quality, but the goal is often to gather as many answers as possible. The target number of responses can be achieved by following a few steps before posting the questionnaire. The questionnaire should start with a simple introduction, explaining the purpose and with questions that awakes the interest of the respondents. Layout and the length of questionnaire are also believed to have significant impact on the response rate. The type of questions that should be avoided, if possible, are open questions where the respondent is required to fill in an answer without any answering options. (Bryman & Bell, 2011, 79-80)

The questions in the questionnaire will be partly the same as those used by Sabita Mahapatra and Abhishek Mishra used in their study "Acceptance and forwarding of electronic word of mouth" (2017). The questions are divided into four different categories and the response options are: 1 (strongly disagree) - 7 (strongly agree). I felt it was important to ask similar questions as in the previous study in order to be able to compare the results properly. The questions I will use in the questionnaire are listed here below. (Bryman & Bell, 2011, 237)

Message credibility

1. I think the review is credible
2. I think the review is trustworthy
3. I think the review is convincing
4. I think the review is honest
5. I think the review is questionable
6. I think the review is authentic
7. I think the review is reasonable

Source credibility

1. I consider online reviewer as being sincere
2. I consider online reviewer as being honest
3. I consider online reviewer as being trustworthy
4. I consider online reviewer as being credible
5. I consider online reviewer as being biased
6. I consider online reviewer as being reputable
7. I consider online reviewer as being truthful

Tie strength

1. Likelihood of sharing personal confidences with peers
2. Likelihood of spending some free time socializing with peers
3. Likelihood of performing a large favour for your peers
4. Likelihood of peers performing a large favour for you

AEWOM

1. I closely follow the suggestions of the comments on the website
2. I agree with the opinion suggested on the website
3. I am likely to accept the comments on the website
4. I am influenced by the comment on the website when making a decision

3.1.3 Respondents

The questionnaire will be shared on Facebook, which means that most of the respondents will be from my own social network. Many of the potential respondents live in Finland and most of them are between 25-30 years old. But also, my friends and parents will help collecting answers by sharing the questionnaire with their own network, which widens the group of respondents a bit. I especially expect several answers from my parents' network, which means respondents who are between 50-60 years old. If the respondents represent different age groups, the answers will give more of a general view of how Finnish adults think about the use of referrals.

To increase the data sample, I have decided to conduct a product lottery in hope to awake interest among potential respondents who would otherwise not consider answering the questionnaire. According to a study by Mette Aadahl & Torben Jorgenen (2003) a lottery can affect positively on the number of respondents and is therefore something to consider. The authors researched the difference between two different groups where only one group was offered the possibility to participate in a lottery. The group that was offered the possibility to participate in the lottery had a response rate of 63.4% while the response rate for the other group was 60.4%. So even if the difference is not significant it can still have an impact on the response rate and is therefore worth implementing. (Aadahl & Jorgenen, 2003, 941)

3.2 Method

This chapter provides information on the most important factors to consider when conducting quantitative research. The information helps in the analysis of the data, which ultimately facilitates the understanding of the results of the study.

3.2.1 Multiple regression analysis

Multiple regression (R2) is a correlation involving the relationship between a continuous predictor and a continuous outcome. The continuous predictor is usually explained by the time taken to complete a test, while the continuous outcome is usually related to exam grades. Multiple regression allows for more complicated analyses to be able to understand the correlation between predictors and continuous outcomes. MRAs always consist of two or more independent variables. (Knapp, 2018 p. 4–5)

Grading of an exam is a typical example of how multiple regression can be used for data analysis. Rather than just looking at the time students spend on an exam, it could also be that age, gender, and academic year influence the expectations. Multiple regression identifies the variable most strongly correlated with grade (p □ .05). It also gives the percentage value for all predictor variables and finally it ignores the variables that are not statistically significant ($p >$ .05), which means that they are not considered in predicting the outcome. The predictor variable is the one with greatest influence on the outcome. (Knapp, 2018, 5–6)

The mathematic formula (Djurfeldt et al., 2003, 350):

$$y = a + b_1x_1 + b_2x_2 + \ldots + b_kx_k$$

y = dependent variabel
x = quantitative or binary independent variable

The F-test is an important part of the MRA in determine whether a significant part of the variance of the dependent variable (y) is included in the regression. In addition, the coefficient R2 or adjusted (adj.) R2 expresses the extent to which the total variance of the dependent variable can be inferred from the independent variables. Once it has been confirmed that the regression performed is statistically significant, the analysis of the t-test for the beta coefficients is initiated to determine whether there is a statistically significant relationship between the independent variables (x) and the dependent variable (y).

After deciding whether the beta coefficients are significant, one usually looks at the beta coefficient that describes the direction of the relationship. If the beta coefficient takes a positive (negative) value, the relationship is positive (negative). In the dissertation, the analysis of variance ANOVA is applied to evaluate the reliability of the chosen regression model. The reliability of the beta coefficients of the model is ensured by multicollinearity tests. The Pearson coefficient was also measured to show that the independent variables are not correlated strongly with each other. (Djurfeldt et al., 2003, 350-351)

3.2.2 Reliability & validity

Reliability refers to the reliability of measuring instruments and measurement results used in a study (Bryman & Bell 2011, 170). Reliability often occurs in the context of quantitative studies. An important question regarding the reliability of a quantitative study is whether the results would be the same if the study were conducted again or whether they would be subject to the influence of chance (Bryman & Bell 2011, 62-63).

Reliability means that there is more than one item to measure, that several respondents answer the same questions and that an overall result should be formed. The different indicators should not refer to different factors, so it must be ensured that they are related to each other. Otherwise, there is a possibility that the items indicate something they are not supposed to indicate. This internal reliability can be tested using the split-half method. In this method, all indicators are split into two different groups and the degree of correlation is calculated to see if the results differ from each other. The calculation results in a number called a coefficient, which can be either 0 (no correlation or internal consistency) or 1 (correlation and internal consistency detected). (Bryman & Bell, 2011, 158)

Validity is according to Bryman & Bell (2011, 157) the most important criterion of a research. Validity of a test can show whether a specific measure does reflect to the concept that it is supposed to. The authors mention that the assessment of measurement validity requires that the research is reliable and that is why validity and reliability have such a strong connection. Validity can be divided into four different categories: measurement, ecological internal and external, of which internal and external are usually more critical. *Measurement validity* is usually mentioned in quantitative research and typically referred as construct validity. It has to do with questions that dispute whether the measurement is really accurate and whether it measures what it is supposed to. “Does the IQ test really measure variations in intelligence?”. (Bryman & Bell, 2011, 43)

Ecological validity is concerned whether the social scientific findings are taking in account the normal everyday life and social settings. These can be for example opinions, values, and general life conditions. This in an important topic since different research is constantly made for business or school purpose and even if most findings are technically valid, they might still not have anything to do with the factors concerning a normal everyday life. (Bryman & Bell, 2011, 43)

Internal validity mainly deals with causality and is therefore concerned about the causal relationship between two or more variables. If for example x causes y, how can we still be sure that x is responsible for the variation in y and not something else. Internal validity questions how we can be so sure that magnitude of consequences does cause some variation in general awareness. When discussing the issues of causality, then usually independent variable is referred to as the one having the clausal impact and dependent variable as the effect. (Bryman & Bell, 2011, 43)

External validity mainly questions whether the study results can be generalized beyond the research context. How people or organisations are being chosen to participate in research is crucial regarding external validity and it is also one reason why so many choose to generate representative samples. Some authors have tried to apply the concepts’ reliability and validity to the practice of qualitative research, but others argue that the basis of these ideas in quantitative research renders them inapplicable to or inappropriate for qualitative research. (Bryman & Bell, 2011, 43-44)

3.4 Substantial coefficients and variables

The substantial coefficients and variables that are used in the study will be presented and described in the following chapter. The chapter begins with and description of the coefficient R^2 and adjusted (adj.) R^2. The coefficient adj. R^2 is used to describe the total explanatory power of the dependent variable (Beisland 2009, 11). After a review of R^2 follows a description of the beta coefficient and the p-value. Finally defined the independent variables Source credibility, Message credibility and Tie strength, as well as the dependent variable AEWOM.

3.4.1 R^2 and adj. R^2

The square of the multiple correlation coefficient is R^2 and it indicates how the dependent variable is explained by the independent variables. The R^2 tends to increase when more variables are added to the equation. The number of independent variables can make the interpretation very difficult, even if the equation gives an excellent explanation. Therefore, it is important to find a balance between the factors that make the explanation as easy as possible, without forgetting about the usefulness of the actual results. (Beisland, 2009, 11)

The coefficient R^2 is often used in studies where the intention is to measure the development of value relevance during a given period and between different samples. It is also common that the adj. R^2 coefficient is used in situations where the goal is to adjust the effect of the scarce number of observations in the regression analysis. (Beisland, 2009, 11)

The usage of R^2 coefficient divides a lot of opinions both for and against. Djurfeldt et. al (2003, 336) argue that adj. R^2 is a better measure when the sample consists of less than 200 observations. The idea in MRA is that the more the variance in the dependent variable can be derived from the independent variables the stronger and more reliable is the relationship between these two. (Djurfeldt et al., 2003, 334). There have also been arguments where it is stated that R^2 is not used correctly to measure the relation between dependent- and independent variables. According to Garry King (1986, 678) there are other statistics that can measure the relation more accurate and answer theoretical questions that R^2 is not capable of.

3.4.2 Beta coefficient

In relation to regression, the beta coefficient expresses the slope of the regression line. Whether the relationship between an independent and dependent variable is positive or negative can be judged based on the beta coefficient. A positive value indicates that the relationship is positive, while a negative value indicates that the relationship is negative. As the beta coefficient assumes a negative value and the independent variable increases, means that the dependent variable decreases. To be able to have an existing relationship between an independent and dependent variable, the beta coefficient cannot assume the value zero. (Djurfeldt et al., 2003, 165-166).

3.4.3 P-value

In connection with hypothesis tests, the p-value is most often used to assess whether the null hypothesis can be rejected. The value replaces the z-value that is used a lot in random sampling (Djurfeldt et al., 2003, 195) In connection with hypothesis testing, it is important to ensure that a hypothesis is not exposed to too much influence by chance. Djurfeldt et al. (2003, 193) states that null hypothesis testing is performed to calculate the probability of coincidence to have an impact and connection in a regression. Based on the p-value can be determined whether the null hypothesis can be rejected. When p-value is 0.05 (5%) or below, a relationship can be considered statistically significant (Djurfeldt et al., 2003, 196).

4 Analysis & Results

In this chapter I will go through the descriptive statistics for data that has been used, and I will also present the results of the study. Finally, I will answer the hypotheses that have been set for the study.

Of the total sample of 102 respondents, 51.9% were women and 85.3% answered the questionnaire in Finnish instead of English. The answers were received quite evenly from different age groups. Most answers came from respondents 21-30 years old (38.4%), 31-40 years old (17.4%) and 51-60 years old (17.4%). When asking about how often the respondents read or react to referrals, the answers were again evenly divided between different answering options. Most respondents read referrals every month (36%). The other options were: Every week (27.9%), less than every month (24.4%), never (9.3%) and every day (2.3%).

4.1 Descriptive statistics

The main purpose of descriptive statistics is to summarise the data collected in a few parameters so that it can be analysed more easily. Descriptive statistics are used to understand the nature of the data set collected, no matter how large the data. Figures and charts are usually used as tools to summarise the data in a simpler format. Continuous data is usually presented in different ways, but the most common way is to use a histogram. A histogram is a bar chart that shows the number of responses for each alternative. Continuous variables are usually summarised with nine different statistics: Count, Mean, Median, Mode, Standard Deviation, Variance, Minimum & Maximum and Range. (Knapp, 2018, 4)

Number shows how many elements there are in a sample, while *mean* is the statistical name for average, meaning the average number when adding up all numbers and dividing them with as many numbers there are to choose between. *Median* stands for the middle value, which means that half the values should be below the median value and the other half should be above it. *Mode* is the most common number that occurs in the data set. It is possible that a data set has several modes, which means it would be referred as a multimodal variable. (Knapp, 2018, 5-7)

Standard deviation shows the dispersion of the numbers within the variable. This means that a data set that includes mostly same numbers will have a lower standard deviation, than for example a data set where the variation between numbers is higher, making the standard deviation also higher. *Variance* is simply the standard deviation squared. This is not usually included in a report, but more used as information among other statistics. *Minimum & Maximum* describes the lowest and highest numbers in the data set. *Range* means the highest number of the data set minus the lowest number. (Knapp, 2018, 7-9)

		MC	SC	TS	AEW
N	Valid	714	714	408	408
	Missing	0	0	306	306
Mean		4,3361	4,2409	4,6029	4,3284
Median		4,0000	4,0000	5,0000	5,0000
Std. Deviation		1,34994	1,32648	1,42432	1,37366

Variance	1,822	1,760	2,029	1,887
Skewness	-,391	-,243	-,630	-,424
Std. Error of Skewness	,091	,091	,121	,121
Kurtosis	-,302	-,509	-,022	-,177
Std. Error of Kurtosis	,183	,183	,241	,241
Minimum	1,00	1,00	1,00	1,00
Maximum	7,00	7,00	7,00	7,00

Chart 1: Descriptive data

The descriptive data helps to understand the characteristics of data set which is one the most important parts of a statistical data analysis

In my analysis I decided to diverge a bit from the most common statistics listed by Knapp (2018) and instead gather the statistics that I considered to be the most interesting for my study. In the chart above can be seen that the number between "valid" and "missing values" is quite significant. TS (Tie Strength) and AEW (AEWOM / Acceptance of electronic word-of-mouth) have several answers missing because of the answers from each category were combined to obtain an average answer for all similar questions. These categories had fewer questions, which is why the statistics show them as missing values.

Most of the statistics that I received from my data set is similar between the different variables. These will be explained more specifically when showing the histograms. I decided to add Skewness and Kurtosis to the statistics, because I found that information interesting and important for my study. *Skewness* and *kurtosis* describe how oblique respectively peak the distribution is. A negative value for skewness indicates that the distribution is sloping more to the right, while a positive value indicates that the distribution is sloping to the left. A symmetrical distribution (normal distribution) assumes the value zero (Djurfeldt et al., 2003, 55). Kurtosis indicates the peak or extension of the distribution. According to Djurfeldt et al. (2003, 55), a negative value means that the distribution is flat, while a positive value shows the distribution to be peaked. In summary, the distribution is skewed on the x-axis, while a peak distribution is seen on the y-axis.

Figure 6: Message credibility

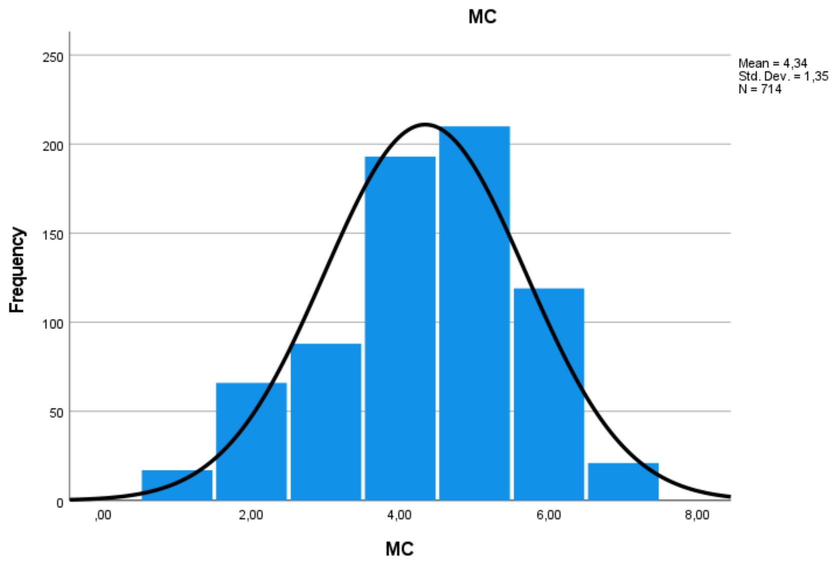
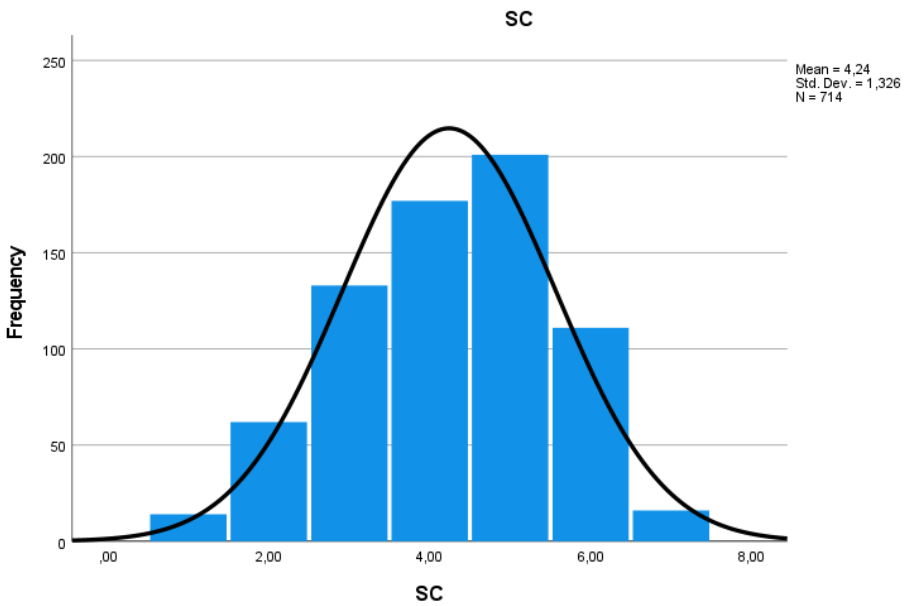


Figure 7: Source credibility



The questionnaire included seven questions for both message credibility and source credibility and there were 102 responses, which explains why there are 714 answers in total. The average answer has clearly been in the middle of the scale (MC 4,34 & SC 4,24) and this shows that most of the respondents have a more positive than negative view on online referrals and towards reviewers who submit the referrals. Standard deviation (MC 1,35 & SC 1,326) and variance (MC 1,822 & SC 1,760) are also quite low for both, which already indicates on many similar answers between the respondents.

The skewness (MC -0,391 & SC -0,243) and kurtosis (MC -0,302 & SC -0,509) are negative for both independent variables and that indicates on answers sloping more to the right. The negative kurtosis shows that the distribution is flat. There are also no significant differences in the total answers between these two variables.

Figure 8: Tie Strength

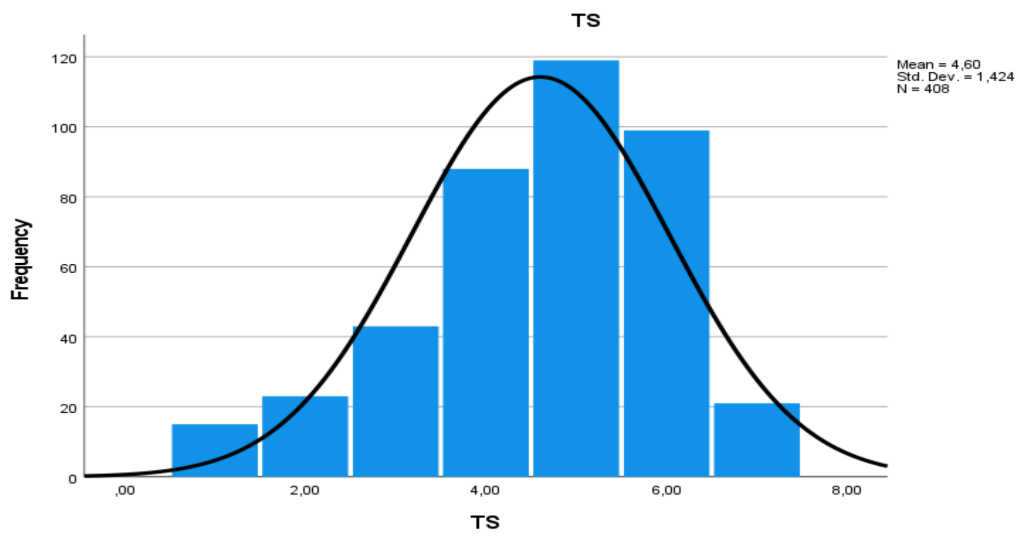
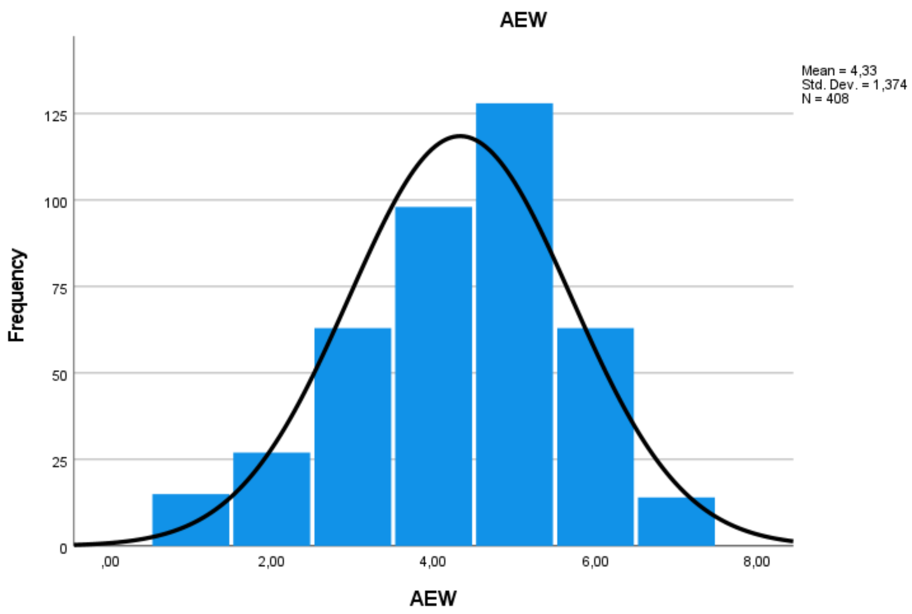


Figure 9: AEWOM



The independent variable Tie strength and dependent variable AEWOM, included both four questions, which explains why both categories have 408 answers in total. The average answers (TS 4,60 & AEW 4,33) for these two categories lean more to the positive side, indicating that the respondents have strong ties with their social network and that they are more likely to accept the referrals that they read about. Standard deviation (ST 1,42 & AEW 1,37) and variance (ST 2,02 & AEW 1,88) is low for both, but especially for Strong ties can be seen that the answers are divided a little more to different answering options, unlike for the other categories. The skewness (ST -0,630 & AEW -0,424) and kurtosis (ST -0,022& AEW -0,177) are negative for both variables and that indicates on answers sloping more to the right. The negative kurtosis shows that the distribution is flat, even if Tie strength is very close to a positive value, which would mean closer to an optimal distribution curve.

4.2 Perceived task value

Perceived task value is often measured when data has been collected by a questionnaire. The measurement is also called Cronbach's alpha (CA) and its purpose is to understand the internal consistency of a questionnaire that includes multiple scales and items. It is one of the most common ways to measure reliability between the questions. When the measurement is parallel, it means that there are equal variances and covariances and if the measurements are congeneric, then the results can be presented with uncorrelated errors. (Bonett, 2014, 1)

CA can vary depending on the ample size, because a small sample can make the result look impressive, but a high confidence interval could still have a low limit indicating on poor reliability. The sample size and its effect on CA can be experienced as worrying, because a limited sample has a bigger possibility to be below 7, which indicates on low internal consistency and that alone can be enough for a manuscript to be rejected. However, there are no minimal acceptance value and therefore the reliability of the entire study is not only depending on this value. (Bonett, 2014, 2)

	Reliability Statistics		
	Cronbach's Alpha	Cronbach's Alpha Based on Standardized Items	N of Items
MC	,828	,823	7
SC	,857	,837	7
TS	,911	,913	4
AEWOM	,905	,906	4

Chart 2: Cronbach's alpha

The reliability statistics indicate the CA coefficient where the target is to score over .7 to have internal consistency.

According to the chart above the coefficient $\alpha = .828$, indicates on higher value than required, which means that the questions related to message credibility are reliable. CA for source credibility indicates that it can be considered reliable with a value of 0,857, which is higher than for message credibility. Tie strength has the highest CA with a value of 0,911. Meaning that it is well above the required limit to be considered reliable. Like the others, AEWOM is clearly above the required limit with a value of .905. According to the above chart, all variables show a high internal consistency.

Item statistics offer important value regarding the actual questions in the questionnaire. The table shows if there are any specific questions that are not in line with the other questions and in this kind of situation the questions are usually removed to retain the reliability of the questionnaire.

Item Statistics			
	Mean	**Std. Deviation**	**N**
MC1	4,3824	1,29005	102
MC2	4,3235	1,38708	102
MC3	4,3824	1,40045	102
MC4	4,4216	1,41726	102
MC5	3,9902	1,30114	102
MC6	4,3431	1,29351	102
MC7	4,5098	1,33309	102
SC1	4,0294	1,34574	102
SC2	4,4020	1,39490	102
SC3	4,1863	1,44009	102
SC4	4,3824	1,45591	102
SC5	4,1765	1,12057	102
SC6	4,2059	1,18854	102
SC7	4,3039	1,30338	102
TS1	4,2843	1,48509	102
TS2	4,8627	1,50273	102
TS3	4,6569	1,36791	102
TS4	4,6078	1,29092	102
AW1	4,1863	1,41932	102
AW2	4,1176	1,31476	102
AW3	4,3235	1,31376	102
AW4	4,6863	1,39291	102

Chart 3: Item statistics

The chart shows that the questions have similar values under "mean" and this indicates that they are in line with each other. MC5 is the only question that is below 4, but there is still no significant difference compared to the other results, which is why there is no need to remove the question to retain the reliability.

4.3 Hypotheses and research questions

The regression coefficients constitute an important part of the regression equation, and it should always be analysed together with hypotheses to see whether the direction of the effects meets the expectations. (Huizingh, 2007, 14)

Coefficients

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	,849	,235		3,607	<,001
	MC	,277	,056	,276	4,925	<,001
	SC	,222	,055	,229	4,082	<,001
	TS	,287	,039	,298	7,339	<,001

Chart 4: Relationship between dependent and independent variables

H1, H2 and H3 are assuming that source credibility, message credibility and tie strength are positively associated with AEWOM. The hypotheses can be stated as correct because there is a positive relationship between the independent variables and the dependent variable. The beta-coefficients are also statistically significant at a 1% level (<0.01).

Unstandardized B shows the relationship between the independent variables and dependent variable. Tie strength (28.7 %) is according to the results affecting the acceptance of electronic word-of-mouth more than source credibility (22.2 %) and message credibility (27.7 %). However, all the variables are close to each other and therefore there cannot be claimed to be a clear difference between them.

Coefficients

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	,607[a]	,369	,364	1,09551

Chart 5: Dependent variable explained by independent variables

Adjusted R square indicates to what extent the dependent variable can be explained by the independent variables. According to the chart above, the proportion of variance is 36.4%. The residual variance that cannot be explained by the independent variables is 63.6%.

4.4 Collinearity diagnostics

In studies with multiple regression analysis, it is advisable to check for the existence of any inaccuracies in terms of mutual correlation between two or several independent variables. Multicollinearity distorts the beta coefficients for the independent variables, which impairs the reliability of the results in a study. (Djurfeldt et al., 2003, 387-388)

Collinearity diagnostics measures multicollinearity, which refers to correlation among independent variables. SPSS offers two ways to measure this, namely tolerance and VIF. SPSS treats one the independent variables as an independent variable to compute the R^2 of the regression equation. The tolerance is calculated by counting one minus R^2 and the results reflect that part of a variance that cannot be explained by the other independent variables. When the tolerance is low, it means that there is some multicollinearity detected. The critical minimum value for Tolerance is 0.40. Anything below 0.40 suggests on multicollinearity. VIF stands for Variance Inflation Factor and is calculated by dividing number one with the tolerance. A rule number is that VIF values above 10,00 are considered to indicate multicollinearity. (Huizingh, 2007, 5)

VIF values from 5 above can already indicate on some problems with multicollinearity. In situations where VIF value for some reason exceeds 10,00, indicates it a significant problem for the multicollinearity test. These situations are usually solved by removing the variable that is causing the problem and by repeating the analysis. Multicollinearity can also be solved by increasing the sample size, which can be an easier way to approach the problem in some studies compared to removing a variable completely. (Knapp, 2018, 7)

Coefficients

Model	Understandardized Coefficients B	Coefficients Std. Error	Standardized Coefficients Beta	t	Sig.	Collinearity tolerance	Statistics VIF
(Constant)	0,849	0,235		3,607	<0,001		
MC	0,277	0,056	0,276	4,925	<0,001	0,497	2,011
SC	0,222	0,055	0,229	4,082	<0,001	0,497	2,012
TS	0,287	0,039	0,298	7,339	<0,001	0,950	1,053

Chart 6: Multicollinearity

The chart above indicates that there is no multicollinearity that would distort the interpretation of the beta coefficients for the independent variables. Tolerance value for MC (0,497) and SC (0,497) are slightly above the critical minimum, while TS is clearly above the limit (0,950). VIF values for MC (2,011), SC (2,012) and TS (1,053) are all below the maximum of 10,00 and even below 5,00, which is also a limit that can indicate on problems regarding multicollinearity.

4.4 Analysis of Variance

The ANOVA test (Analysis of Variance) is like t test and is usually used to compare three or more groups to each other, while t test is commonly used for comparing two groups at same time. (Knapp, 2018, 4) ANOVA can be used for testing whether means of different groups are equal. ANOVA can be used for measuring for example if tennis players pay equal amounts on tennis court rental or are there significant differences between players who play singles, doubles, and both. Analysis of variance can also be used for measuring if tennis players from three different age groups play equally often every month. (Huizingh, 2007, 3)

Analysis of variance can be divided in two different groups: One-Way ANOVA and Multiple-factor analysis of variance. *One-Way ANOVA* means that the grouping has only one variable. If the goal is to measure the differences among tennis players, then the groups must be divided either by age or gender, for example young, adult and veteran. *Multiple-factor analysis of variance* differs from One-way ANOVA with the number of variables in the groups. A similar example with tennis players would look different with multiple-factor analysis, because the groups could be divided to different groups where the age and gender are combined for example young-male, young-female, and adult-male. (Knapp, 2018, 2-3)

According to Djurfeldt et al (2003, 281), the critical F-value at 5% significance level is 4.60, while at the 1% significance level it amounts to 8.86. The F-value is calculated by dividing the mean squared deviations (mean square) of the independent variables with mean square for the residuals. Mean square is calculated by dividing the sum of squares by the number of degrees of freedom (df) (Djurfeldt et al., 2003, 337)

The mathematic formula for calculating F-test: (Huizingh, 2007, 10)

$$\text{F-value} = \frac{\text{Mean sum of square effect}}{\text{Mean sum of squares residual}}$$

ANOVA

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	283,128	3	94,376	78,637	<,001[b]
	Residual	484,862	404	1,200		
	Total	767,990	407			

Chart 7: Analysis of variance

The regression model is significant at the 1% level (p-value <0.01). The F-value 78.637 exceeds the value 8,86. A significant part (283,128) of the total variation in the dependent variable can be derived from the independent variables, the remainder (484,862) from the residuals cannot be identified by the model

4.5 Pearson correlation coefficient

The bivariate Pearson correlation coefficient is used for correlation analysis, which indicates whether there is a significant relationship between two variables. The correlation analysis shows both strengths and weaknesses between variables. The correlation answers questions such as whether tennis players who play regularly spend more money on court rent than players who play less often. The analysis can also be used to measure the relationship between age and expenditure on tennis clothing. (Huizingh, 2007, 2)

The correlation assumes that the variables are either ratio or interval and the sample must originate from a bivariate normal distribution. When talking about bivariate normal distribution, it usually means that one value of a variable is normally distributed while the other one is not. The correlation cannot reflect to all different types of relationships, but it can reflect the linear relationship. This means that the degree to what two variables reflect to each other is presented with a straight line, meaning if the relationship between two variables is strong but non-linear, then the correlation coefficient is close to zero.

According to Huizing (2007, 2) it can be worth to create a scatterplot before requesting the Pearson correlation coefficient, because the scatterplot can ease the decision whether it is profitable to even request it. However, if the Pearson correlation coefficient is not used, then it should be replaced with other statistics.

Pearson's coefficient can assume values between zero and one. The value zero indicates that there is no connection between the variables, while one indicates a complete relationship. The coefficient can assume positive and negative values depending on whether the relationship is positive or negative (Bryman & Bell 2011, 355). A weak correlation means that the variation of the variables can be explained by other variables not included in the model (Bryman & Bell 2011, 356)

The correlation between two variables is not always positive, which indicates on negative correlation. Knapp (2018, 22) mentions an example where the correlation between time and grade was studied by using the correlation analysis. The result of the study shows that students who use less time on studying ended up scoring higher grades than students who spent more time studying. In this situation the correlation between the two variables (grade and time) is not statistically significant.

Correlations

		AEW	MC	SC	TS
Pearson Correlation	AEW	1,000	,499	,485	,402
	MC	,499	1,000	,706	,207
	SC	,485	,706	1,000	,207
	TS	,402	,207	,207	1,000
Sig. (1-tailed)	AEW	.	<,001	<,001	<,001
	MC	,000	.	,000	,000
	SC	,000	,000	.	,000
	TS	,000	,000	,000	.
N	AEW	408	408	408	408
	MC	408	408	408	408
	SC	408	408	408	408
	TS	408	408	408	408

Chart 8: Pearson correlation coefficient

The chart above indicates that the correlation is statistically significant at the 1% level, and thus one can draw reliable conclusions from the chart. Initially, it can be said that the correlation between the variables is positive. The positive connection is expected and natural to the extent that a positive result contributes to improved acceptance of EWOM.

5. Discussion

This study investigated the difference between three independent variables (message credibility, source credibility and tie strength) and their effect on the dependent variable (AEWOM). A quantitative research method in the form of a questionnaire was used because it was best suited for this type of study. Questionnaire research is used to collect data through a questionnaire or multiple interviews at a single point in time. Otherwise, it would be more difficult to collect quantitative data associated with at least a few variables.

The findings show that all three variables have a quite significant impact on the acceptance of electronic word-of-mouth. Tie strength is according to the results the factor that impacts acceptance the most (28.7 %), before message credibility (27.7 %) and source credibility (22.2 %). When looking at the previous study done by Mahapatra & Mishra (2017), the results show some similarities but also differences compared to the findings of this study. Their results indicate that source credibility (40.2 %) and tie strength (32.1 %) have a significant impact on AEWOM, while the role of message credibility (15.9 %) is less significant.

Strong ties are commonly resulting in higher trust and effective transfer of ideas and information. (Hansen, 1999, 82) The closeness of peers is also increasing the acceptance of EWOM. (Mazzarol et al., 2007, 1486). These factors clearly have a major impact on acceptance and especially when both studies are showing similar results it gives it even more credibility. From a company perspective, this is information that could be used to encourage the current customers to let their friends and different social groups also know about the positive feedback regarding a product or service. It would be interesting to see if the recommendations would be more actively shared to peers if the online sites would for example have a separate "share-button" for the reviews.

Consumer's value and accept recommendations likelier from experts (Willemsen et al., 2012, 508-510). Many of those who submit online recommendations can also be thought as experts, which could explain the significant impact it has on AEWOM, especially in the results of the study by Mahapatra & Mishra (2017).

Message credibility is referred as a review or recommendation that is considered believable and trustworthy. Usually, a message is considered trustworthy when the receiver finds it believable and factual. (Sweeney et al., 2012, 239) This statement can also be agreed on because it clearly has an impact on the results of acceptance. Mahapatra & Mishra (2017) found this part less significant, which they assumed to depend on consumers trust for the message and that they are generally more concerned over source credibility, or they are looking for tie strength as an indicator to accept online reviews. Also, if the recommendation is originating from peers, then the message can be assumed to be credible as well, compared to receiving information for example from a marketeer.

The study indicated on significant internal consistency as measured by Cronbach's alpha. Cronbach's alpha is a measure of internal consistency, which means how closely a group of items is related to each other. It is considered a measure of the reliability of a scale. A "high" value for alpha does not mean that the measurement is unidimensional but is one of the most common ways of measuring inter-question reliability. If the measurement is parallel, it means that the variances and covariances are equal, and if the measurements are congeneric, then the results can be presented with uncorrelated errors. (Bonett, 2014, 1) All the variables indicated a high internal consistency and values well above the required limit of .7. The item statistics also provided crucial information about the questions used in the questionnaire. According to the statistics, all questions were consistent with each other, and no questions had to be removed to maintain the reliability of the questionnaire.

The questions of the questionnaire had already been used in a previous study with some good results, which is why there was no doubts over internal consistency or the reliability of the questionnaire. The questions were in general very good and afterwards, it is easy to see that they really helped to obtain the needed information for the study. A qualitative research method could also be used instead of the quantitative method if the goal would be to obtain more precise information from a limited sample. However, for this study, the quantitative method worked well and reached all the expectations.

Overall, I am pleased with the results of the study, and I also believe that the comparison between the two studies gives them both more credibility because the results are similar but still there are some details that differ from each other, making it interesting to compare. Mahapatra & Mishra (2017) had a sample of 324 respondents from India, while this study consisted of 102 respondents from Finland. It can be assumed that consumers in both countries share similar opinions regarding digital referrals and the acceptance of it. The Finns might also be more concerned about the message credibility while consumers in India have more trust for the message if the source is valid, or if tie strength is included as an indicator. However, due to the limited size of the sample I would not go as far as claiming significant differences between these two populations.

6. Conclusion

The purpose of the study was to research consumers' attitudes towards digital referrals by using the AEWOM model. The purpose has been researched by using four different research questions and three hypotheses. The research questions that were used in the study were "what are referrals in marketing?", "How and why do consumers rely on referrals?", "What factors influences acceptance of electronic word of mouth?" and "To what extent can the variance in acceptance of electronic word of mouth be explained by these factors? (Which of these factors is the most influential?)". The first and second research questions were answered in the literature review, while the other two questions were answered together with the hypotheses by using multiple regression analysis. The research questions supported the study and gave it a clear guideline to follow.

The findings of the study show that there are different kinds of traditional and digital referrals that can impact the consumers behavior. WOM is likely to be the most common traditional referral and something companies have relied their marketing majorly on. Consumers still use and rely on recommendations they receive in form of WOM, but with digitalization a significant part of WOM has moved online, creating EWOM. Influencers, recommendation systems and EWOM are the most common modern referrals from which consumers receive information about products and services. Consumers' attitudes towards the usage of referrals are quite evenly divided both for and against. Many consumers feel that referrals in general help them save time and decide by narrowing the options, while others want to keep their autonomy in decision-making by avoiding any external manipulation which for example could lead to impulse purchases.

The hypotheses that were used in the study assumed that message credibility, source credibility and tie strength have a positive impact on consumers' acceptance of digital referrals. To answer the hypotheses a quantitative study was implemented together with a multiple regression analysis. The independent variables that were used for the analysis were message credibility, source credibility and tie strength, while AEWOM was used as the dependent variable. The results show that all independent variables have a positive impact on the dependent variable. Tie strength was according to the results affecting AEWOM more than source credibility and message credibility. However, all the variables were close to each other and therefore there cannot be claimed to be a clear difference between them.

To assess the reliability of the study, a variance test ANOVA was conducted to illustrate how the total variation in the dependent variable is distributed among the independent variables in the regression model. Pearson correlation test was also conducted between the independent variables to show that the independent variables are not overly influenced by each other. Since the study involves a multiple regression analysis, a multicollinearity test was performed to ensure that there is no mutual collinearity between the independent variables. The results were significant in the sense that the beta coefficients for the independent variables could be analysed, and value relevance assessed.

The value relevance of the result was measured with adjusted R^2 coefficient and the value relevance for the variables was measured with beta-coefficient. The reliability and validity were on a good level, and this was proven by testing multicollinearity, variance analysis and internal collinearity. The internal validity was also on a good level because the relationship between the independent and dependent variable was confirmed by ascertaining that the beta coefficients are statistically significant at the 1% level. Also, to show that a significant part of the total variation in the dependent variable is caused by independence variables, a variance analysis was performed for all samples. The results should only be generalised to adults in Finland, since most of the respondents had chosen to answer in Finnish and it can be assumed that the minority that answered in English are Swedish-speaking Finns. Based on these factors and the discussion I conclude that the reliability and validity are on a good level.

This study did not investigate the differences in acceptance of EWOM between men and woman or differences between different age groups. This could be an interesting topic for future research since it is possible that the differences would be significant, especially between consumers from different generations. Another interesting topic for future research could be to investigate how consumers' acceptance towards referrals have changed when having moved from WOM to EWOM.

References

Aadahl, M. & Jorgensen, T. (2003) *The effect of conducting a lottery on questionnaire response rates: A randomised controlled trial.* European Journal of Epidemiology. Vol. 18, No. 10, pp. 94

Aaltonen, M. (2019) *Tekoäly – Ihminen ja kone*. Helsinki: Alma Talent. pp. 94–161

Alboqami, H., Al-Karaghouli, W., Baeshen, Y., Erkan, I., Evans, C. & Ghoneim, A. (2015) *Electronic word of mouth in social media: the common characteristics of retweeted and favourited marketer-generated content posted on Twitter.* Internet Marketing and Advertising. Vol. 9, No. 4, pp. 345-346

Alrasheedi, N., Hassan, S. & Leen, J. (2021) *A Conceptual Model Development for Persuasive E-Wom Reviews: Challenges and Future Prospect*. University of Sains Malaysia. pp. 2124-2128

AMA. (n.d.) *What is digital marketing?* Retrieved 08.03.2021 from: https://www.ama.org/pages/what-is-digital-marketing/

Arbatskaya, M. & Konishi, H. (2016) *Consumer referrals*. International Journal of Industrial Organization. Vol. 48, pp. 35-36

Bagozzi, R. & Warshaw, P. (1990) *Trying to Consume*. Journal of Consumer Research. Vol. 17, No. 2, pp. 131

Bansal, S. & Baliyan, N. (2019) *A Study of Recent Recommender System Techniques.* International Journal of Knowledge and Systems Science. Vol. 10, No. 2, pp. 30-45

Behera, R., Gunasekaran, A., Gupta, S., Kamboj, S. & Bala, P. (2020) *Personalized digital marketing recommender engine*. Vol. 53, pp 1-15.

Bjørlo, L., Moen, Ø. & Pasquine, M. (2021) *The Role of Consumer Autonomy in Developing Sustainable AI: A Conceptual Framework*. Department of Industrial Economics and Technology Management, Norwegian University of Science and Technology (NTNU) Trondheim, Norway. pp. 11-12

Bonett, D. (2014) *Cronbach's alpha reliability: Interval estimation, hypothesis testing, and sample size planning*. Journal of Organizational Behavior. pp. 1-2

Botha, M. (2018) *The 15 most important AI companies in the world.* Retrieved 15.1.2021 from:

https://towardsdatascience.com/the-15-most-important-ai-companies-in-the-world-79567c594a11

Bortko, K. (2019) *Multi-criteria Evaluation of Recommending Interfaces towards Habituation Reduction and Limited Negative Impact on User Experience*. Procedia Computer Science. pp. 2240-2241

Bradley, S. (2015) *Win the Game of Googleopoly*. New Jersey: John Wiley & Sons, Inc. pp. 1-3

Brown, D. & Fiorella, S. (2013) *Influence Marketing – How to Create, Manage and Measure Brand influencers in Social Media Marketing.* Indianapolis: Que Publishing. pp. 3-7

Bryman, A. & Bell, E. (2011) *Business Research Methods.* New York: Oxford University Press. pp. 54-170

Bughin, J., Doogan, J., & Vetvik, O. (2010) *A new way to measure word-of-mouth marketing*. Retrieved 03.01.2022 from http://designdamage.com/wp-content/uploads/2009/04/A-new-way-to-measure-WOM-marketing.pdf

Cantallops, A. & Salvi, F. (2014) *New consumer behaviour: A review of research on eWOM and hotels.* International Journal of Hospitality Management. Vol. 36, pp. 47-48

Carter, B., Brooks, G., Catalano, F. & Smith, B. (2007) *Digital marketing for dummies*. Chichester: John Wiley & Sons, Inc. pp. 145-146

Chen, S., Qiu, H., Zhao, S., Han, Y., He, W., Siponen, M., Mou, J. & Xiao, H. (2021) *When more is less: The other side of artificial intelligence recommendation*. Journal of Management Science and Engineering. pp. 2-4

Cheung, M., Luo, C., Sia, C. & Huapeng, C. (2009) *Credibility of electronic word-of-mouth: informational and normative determinants of on-line consumer recommendations.* Vol. 13, No. 4, pp. 14-15

Chris, A. (n.d.) *What is digital marketing?* Retrieved 08.03.2021 from: https://www.reliablesoft.net/what-is-digital-marketing/

Dick, S. (2019). *Artificial Intelligence*. Retrieved 01.03.2021 from: https://hdsr.mitpress.mit.edu/pub/0aytgrau/release/2?readingCollection=672db545

Djurfeldt G., Larsson, R. & Stjärnhagen, O. (2003) *Statistisk verktygslåda – samhällsvetenskaplig orsaksanalys med kvantitativa metoder*. Studentlitteratur. Lund. pp. 196–350

Enge, E. (2012) *The art of SEO – Mastering Search Engine Optimization*. Hihgway North: O'reilly Media, Inc. pp. 34-36

Fan, A., Shen, H., Wu, L., Mattila, A. & Bilgihan, A. *Whom do we trust? Cultural differences in consumer responses to online recommendations*. International Journal of Contemporary Hospitality Management. Vol. 30, pp. 150-151

Fayaaz, Z. (2020) *Recommendation Systems: Algorithms, Challenges, Metrics, and Business Opportunities.* Computer Science Department, Wilfried Laurier University, Waterloo, Canada. pp. 4-8

Fesenmeier, D. (2006) *Destination Recommendation Systems Behavioural Foundations and Applications*. London: CAB International. pp. 35-38

GDPR.EU (2021) *What is GDPR, the EU's new data protection law?* Retrieved 11.09.2021 from: https://gdpr.eu/what-is-gdpr/?cn-reloaded=1

Gershgorn, D. (2017) *The Quartz guide to artificial intelligence: What is it, why is it important, and should we be afraid?* Retrieved 25.02.2021 from: https://qz.com/1046350/the-quartz-guide-to-artificial-intelligence-what-is-it-why-is-it-important-and-should-we-be-afraid/

Granovetter, M. (1973) *The strength of weak ties*, American Journal of Sociology, Vol. 78, No. 6, pp. 1360-1380.

Grimsley, S. (n.d.) *What Is the Buying Process in Marketing? - Stages & Overview.* Retrieved: 23.01.2021 from: https://study.com/academy/lesson/what-is-the-buying-process-in-marketing-stages-lesson-quiz.html

Hansen, M.T. (1999*) The search-transfer problem: the role of weak ties in sharing knowledge across organization subunits*. Administrative Science Quarterly. Vol. 44, No. 1, pp. 82

Hardesty, L. (2019) The history of Amazon's recommendation algorithm. Retrieved: 04.01.2022 from https://www.amazon.science/the-history-of-amazons-recommendation-algorithm

Hayes, A. (2021) *Word-of-Mouth Marketing (WOM Marketing)*. Retrieved 14.10.2021 from: https://www.investopedia.com/terms/w/word-of-mouth-marketing.asp

He, S., Hollenbeck, B. & Proserpio, D. (2021) *The Market for Fake reviews.* pp 42-43. Retrieved 09.02.2021 from: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3664992

HG Insights. (n.d.) *Propensity Modelling for Business*. Retrieved 03.03.2021 from: https://datascience.foundation/sciencewhitepaper/propensity-modelling-for-business

Huizingh, E. (2007) *Applied Statistics with SPSS.* London: SAGE Publications Ltd. pp. 2-5

Infosys (2021) *Rethinking retail: Insights from consumers and retailers into an omni-channel shopping experience.* Retrieved 29.08.2021 from: https://www.infosys.com/newsroom/press-releases/Documents/genome-research-report.pdf

Isinkaye, F., Folajimi, Y. & Ojokoh, B. (2015) *Recommendation systems: Principles, methods and evaluation.* Egyptian Informations Journal. Vol. 16, pp. 262

Jalilvand, M., Shekarchizadeh, E. & Samieri, N. (2011) *Electronic word-of-mouth: Challenges and opportunities*. Procedia Computer Science. Vol. 3, pp. 44-45

Kaplan, J. (2016) *Artificial Intelligence: What Everyone Needs to Know*. New York: Oxford University Press. pp. 2–5

Kaushik, G. Prakash, R. (2018) *Collection of data through cookies and smart devices – A case study*. International Journal of Advance Research, Ideas and Innovations in Technology. Vol. 4, No. 5, pp. 458-460

King, G. (1986) *How Not to Lie with Statistics: Avoiding Common Mistakes in Quantitative Political Science*. American Journal of Political Science. Vol. 30, pp. 678

Kembellec, G, Chartron, G. & Saleh, I. (2014) *Recommender Systems.* Hoboken: John Wiley & Sons, Inc. pp. 59-60.

Kent, P (2012). *SEO for dummies.* New Jersey: John Wiley & Sons, Inc. pp. 25-28

Knapp, H. (2018) *Intermediate Statistics Using SPSS*. Los Angeles: SAGE Publications Ltd. pp. 2-22

Komulainen, M. (2018) *Menesty Digimarkkinoinnilla.* Helsinki: Kauppakamari. pp. 150–301

Kulyk, O. (2018) *"This Website Uses Cookies": Users' Perceptions and Reactions to the Cookie Disclaimer.* Technische Universitat Darmstadt, Germany. pp. 2-4

Leskovec, J., Rajaraman, A. & Ullman, J. (2014) *Mining Massive Datasets.* E-book. pp. 336-337. retrieved: 18.12.2021 from http://infolab.stanford.edu/~ullman/mmds/ch9.pdf

Linden, G., Smith, B. & York, J. (2010) *Amazon.com Recommendations – Item-to-Item Collaborative Filtering.* Retrieved 09.10.2021 from: http://citeseerx.ist.psu.edu/viewdoc/download;jsessionid=50BCA380DE5AB838F9A0EFFD888C70F3?doi=10.1.1.221.9347&rep=rep1&type=pdf

Luna-Nevarez, C. & Torres, I. (2015) *Consumer attitudes toward social network advertising*. Journal of Current Issues and Research in Advertising. Vol. 36, No. 1, pp. 10-13

Mahapatra, S. & Mishra, A. (2017) *Acceptance and forwarding of electronic word of mouth*. Marketing Intelligence & Planning. Vol. 35, No. 5, pp. 597-602

Mazzarol, T., Sweeney, J. & Soutar, G.N. (2007) *Conceptualizing word-of-mouth activity, triggers and conditions: an exploratory study*. European Journal of Marketing, Vol. 41, No 11, pp. 1486

McDonald, P., Brannon Donoghue, C. & Havens, T. (2021) *Digital Media Distribution: Portals, Platforms, Pipelines.* NYU Press. pp. 10-15

Mediakix (2022) *The 143 Top Tiktok Influencers – Biggest, Upcoming & Categories*. Retrieved 22.11.2021 from: https://mediakix.com/blog/top-tiktok-influencers/

Milano, S., Taddeo, M. & Floridi, L. (2020) *Recommender systems and their ethical challenges*. AI & Society. Vol. 35, pp. 960-962

Mintz, J. (2005) *Large Firms Increasingly Rely on Employees for Job Referrals*. pp. 1-2. Retrieved 14.1.2021 from: http://unionstats.gsu.edu/8220/Mintz_Large%20Firms%20Rely%20On%20Referrals_2005.pdf

Mitchell, I. (2012) *Third Party Tracking Cookies and Data Privacy*. Northern Kentucky University. pp. 4-8

Núñéz-Valdéz, E., Lovelle, J., Martinéz, O., García-Díaz, V., Ordonez de Pablos, P. & Marín, C. (2012) *Implicit feedback techniques on recommender systems applied to electronic books*. Vol. 28, No. 4, pp. 1186-1193

Paraschakis, D. (2017) *Towards an ethical recommendation framework*. International Conference on Research Challenges in Information Science (RCIS), pp 215

Park, C. & Lee, T. (2009) *Information direction, website reputation and eWOM effect: A moderating role of product type.* Journal of Business Research. Vol. 62, pp. 64-65

Paul, W. & Cara, A. (2015) *The effects of review valence in organic versus sponsored blog sites on perceived credibility, brand attitude, and behavioural intentions*. Marketing Intelligence & Planning, Vol. 33, No. 4, pp. 508-510.

Pelau, C., Niculescu, M. & Stanescu, M. (2020) *Consumers' perception on the advantages and disadvantages of cookies and browsing history.* International Conference on Business Excellence 2020. pp. 830

Press, G. (2017) *Artificial Intelligence (AI) Defined*. Retrieved 09.02.2021 from: https://www.forbes.com/sites/gilpress/2017/08/27/artificial-intelligence-ai-defined/?sh=705678267661

Quentin, A., Carmon, Z., Wertenbroch, K., Crum, A., Frank, D., Goldstein, W., Huber, J., Van Boven, L., Weber, B. & Yang, H. (2017) *Consumer Choice and Autonomy in the Age of Artificial Intelligence and Big Data*. pp. 28-37

Ruan, W-Q., Zhang, S-N., Liu, C-H. & Li, Y-Q. (2020) *A new path for building hotel brand equity: the impacts of technological competence and service innovation implementation through perceived value and trust*. Journal of Hospitality Marketing & Management. Vol, 29, No. 8.

Rubanovitsch, M, D. (2018) *Myyntikapina: Korvaako kone ihmisen myyntityössä?* Espoo: Johtajatiimi. pp. 114-118

Sammis, K., Lincoln, C. & Pomponi, S. (2015) *Influencer Marketing for Dummies*. Hoboken: John Wiley & Sons, Inc. pp. 12-27

Sanders, J. (2016) *Defining Terms: Data, Information and Knowledge*. SAI Computing Conference. pp. 2-4

Schubert, P. (2000) *Virtual Communities of Transaction: The Role of Personalization in Electronic Commerce.* Vol. 10, No. 1, pp. 44-55

Smith, B & Linden, G. (2017) *Two Decades of Recommender Systems at Amazon.com*. IEEE Internet Computing. Vol. 21, No. 3, pp. 12-16

Sterne, J. (2017) *Artificial Intelligence for Marketing: Practical Applications.* New Jersey: John Wiley & Sons, Inc. pp. 21–24

Stinson, C. (2020) *Algorithms are not neutral: Bias in collaborative filtering*. Philosophy Department and School of Computing Queen's University Kingston, ON Canada. pp. 3-8

Swanson, R. & Holton, M (2005) *Research in Organizations: Foundations and Methods in Inquiry*. San Francisco: Berret-Koehler Publishers, Inc. pp. 30-32

Sweeney, J.C., Soutar, G.N. & Mazzarol, T. (2012) *Word of mouth: measuring the power of individual messages*. European Journal of Marketing, Vol. 46, No 1, pp. 237-257.

Tang, T., Winoto, P. (2016) *I should not recommend it to you even if you will like it: the ethics of recommender systems*. New Rev Hypermedia Multimedia. Vol. 22, No. 1, pp. 120

Torres, J., Arroyo, X., Irurita, A. & Solé Moro, M. (2018) *Impact of gender on the acceptance of electronic word-of-mouth (eWOM) information in Spain.* Contaduría y Administración. Vol 63, No. 4, pp. 2

Vuolle, N. (2019) *Tekoälyn hyödyntäminen kuluttajan ostopäätösprosessissa.* University of Jyväskylä. pp. 7–11

Wang, X. & Wei, K. (2006) *Consumers' Acceptance of Electronic Word-of-Mouth Recommendations: Effects of Multiple Communication Elements and Processing Motivation.* International Conference on Information Systems. pp. 282

Ye, Q., Law, L. & Gu, B. (2009) *The impact of online user reviews on hotel room sales*. International Journal of Hospitality Management. Vol. 28, No. 1, pp. 181

Zhang, J., Craciun, G. & Shin, D. (2010) *When does electronic word-of-mouth matter? A study of consumer product reviews*. Journal of Business Research. Vol. 63, No. 12, pp. 1340-1341

Zhang, P., Lee, H-M., Zhao, K. & Shah, V. (2019) *An empirical investigation of eWOM and used video game trading: The moderation effects of product features*. Decision Support Systems. Vol. 123, pp. 9

Zhou, S., Blazquez, M., McCormick, H. & Barnes, L. (2021) *How social media influencers' narrative strategies benefit cultivating influencer marketing: Tackling issues of cultural barriers, commercialised content, and sponsorship disclosure*. Vol. 134, pp. 122

Zoetekouw, K. (2019) *A critical analysis of the negative consequences caused by recommender systems used on social media platforms*. University of Twente. pp. 4-5