Make-up Reviews on YouTube:
The effect of the context of videos on the effectiveness of native advertising

Student name: Nadia Stam
Student number: 410202
Supervisor: dr. R. Jacobs
Date: 21/06/2018

Master Media Studies (Media & Business)
Erasmus School of History, Culture and Communication
Erasmus University Rotterdam

19,281 words

Abstract

In recent years, companies have flocked towards alternative, modern media to advertise and market their products. Social media provides the perfect opportunity to target particular audiences and niche markets, as they often consist of thousands of communities with similar interest. One of the communities that is of a particularly vast size is the beauty community that exists on the video-sharing platform YouTube. Companies compensate YouTubers for talking about their brand by sending them products or sponsoring a video. The main problem that arises with this is that these advertisements can be deceptive in nature, as the audience is often not informed about the existence of the sponsorships. Editorial content like this is often extremely hard to recognize. Many governments and even some YouTubers have acknowledged the possible dangers of this type of native advertising, and urge these creators to more distinctly disclose the sponsorships whenever they can as this could possibly help the consumer to make a more informed decision that is free of deception. However, as consumers can often be resistant towards advertising, a disclosure can also have some negative effects on advertising effectiveness and credibility. An online survey-experiment ($N = 258$) examined whether these assumptions were true. This was done through the use of different titles and description boxes in the same YouTube video, which each presented a different type of disclosure, or lack thereof. The findings of this study indicated that advertising effectiveness was not affected by the contents of the description box, nor by different titles. The additional analyses that were conducted suggested that there is indeed a relationship between advertising effectiveness and sponsorship disclosure, but it is still unclear how this relationship is exactly formed. These findings are in line with the scarce amount of existing literature, which is still inconclusive about the true effects of sponsorship disclosures as different studies state opposite things. Future research should focus on clarifying how exactly sponsorship disclosures, in the various forms that they exist in, affect consumer behaviors and opinions. This could help regulatory bodies, YouTubers and companies alike to establish more ethical ways of advertising, whilst still supporting commercial interests.

Keywords: native advertising, YouTube, sponsorship disclosure, advertising effectiveness, influencers

Table of contents

Abstract & keywords

1. Introduction

In April of 2017, the Dutch Commission of Media published results of research that studies how often brands and product are visible in YouTube videos (Rozendaal & van Reijmersdal, 2017). The reports showed that an increasing number of Dutch YouTubers were promoting products, services and brands through their videos without reporting it to their followers. Since YouTube itself does not have any strict guidelines for this type of advertising on the medium, YouTubers can individually decide to what extent they want to take this paid advertising and promoting of free products. Either, they can disclose to their audience that they were given the products, or they can keep it to themselves and let the viewers find it out on their own terms. The ethics behind this are very hard to regulate and since YouTube is an international medium that operates solely online, it is has been proven difficult to establish any regulations about this. Dutch YouTuber Saske de Schepper read the report and decided that she could not stand the ambiguity of advertising in these YouTube videos, and even more so, the YouTubers that were misleading their followers by deviously talking positively about products that they were given for free or paid to talk about. Together with 17 other Dutch YouTubers, she founded an institution that gives YouTubers guidelines on what they should and should not do when it comes to promoting products or services. This institution, which exists in the form of a website, is called De Social Code (https://www.desocialcode.nl/). For example, if a YouTuber receives a compensation for promoting a product, they should inform the viewer by, for example, showing the cooperation between them and the brand for at least 3 seconds in written text. It would be even better to disclose the sponsorship verbally it in the video. If a product was given for free, or with a discount, YouTubers should disclose this in the description box of their videos. These guidelines were made to help new and experienced YouTubers alike to decrease the amount of deception that younger YouTube viewers experience every day (Meindersma, 2017). Children as young as 6 or 7 are allowed to watch videos of YouTubers filming what they do through the day (vlogging), playing a video game, or pranking their friends. These younger audiences are less perceptive of and resistant to native advertising that often appears in these videos than somewhat older audiences. The question is, however, whether these guidelines are effective in reaching the goals that were set. For instance, are viewers of YouTube videos indeed more perceptive of

advertising if it is shown in the description box or title? In this research, it is aimed to provide an answer to this question by answering the following research question:

How does the context of YouTube videos affect the effectiveness of native advertising?

Ever since its establishment in 2005, YouTube has proven itself to be the perfect platform for consumers to search for content that can satisfy each and every individual. In 2017, a billion hours of videos was watched every single day (Press - YouTube, 2018). This number holds all of the different types of videos that exist, from music videos to game reviews, and from news items to movie trailers. Audiences flock to the website in search for entertainment, information and communities; people with similar interests watch similar videos and can comment on these, creating a type of mini-chat under each one. One of the consumer industries that has the biggest community on YouTube, is the beauty industry (García-Rapp & Roca-Cuberes, 2017). Thousands of girls (and boys, for that manner) post videos about their make-up routines, their favorite perfume and how to best style your hair in a ponytail. This community poses an vast opportunity for companies in the field to promote their products (Chang, Molesworth & Grigore, 2015). Even though typical advertisements before these videos may seem as the most logical and easy way to do so, cosmetic companies have started to develop a more sophisticated form of marketing and brand integration. Once beauty-YouTubers has a certain amount of subscribers to their channel, these companies will start sending them products and asking the YouTuber to review them. Paying people to given reviews about the quality of certain products and services has been an important marketing technique that businesses use for decades, which originated with the reviewing of restaurants and movies (Eliashberg & Shugan, 1997). However, the main difference between these critics and the YouTuber is that the YouTuber is independent, whereas the traditional critic is bound to publish their review in a newspaper or magazine, which thus limits them in their freedom in what to say. Either way, both traditional and modern reviewing processes that are based on companies providing the product and/or service to the reviewer have commercial interests. In the end, the company sends their product because they want their target audience to hear that their product is good. Additionally, they want it to come from of an opinion leader that has a particular influence towards the target audience. There are various ways in which a YouTuber can go about reviewing a product. Sometimes, the YouTuber will get paid to talk about the product, or they will get a coupon code to give to their viewers for which they get

2

a commission for every time the code is used. For example, YouTuber 'Jane Doe' may mention her discount code JANE20 in her video for 20% off on products on a certain brand's website. For every time someone uses this code on the website, Jane gets a certain amount of money from the company. And other times, the YouTuber gets to decide for themselves whether they wish to mention the product, and they do not get compensated for it at all.

The persuasion of consumers is and has already for a long time been a heavily debated topic, especially within the scope of ethics related to advertising and marketing communication (Ikonen, Luoma-aho & Bowen, 2017). When a YouTuber posts a video where he/she positively reviews a product and it turns out he/she gets paid for it, this could be interpreted as wrongful persuasion. Communities on YouTube are often very close-knit, and viewers take the advice of their favorite YouTubers into account when they are going through the process of buying a product. If a YouTuber abuses this trust by getting paid to recommend it, this can be considered unethical. That is why, as mentioned before with the story of 'The Social Code', a lot of YouTubers have lately started to either mention that they are being sponsored for the review, or they disclose this in their description box. However, this may in turn affect the effectiveness of the product promotion. If a viewer knows that a YouTuber is being paid to review the product, does he/she still wants to buy it? These kinds of questions are exactly the reasons that make YouTube marketing an important and interesting topic to study.The key concept in this study is the disclosure of sponsorships. Advertisers are getting more creative in their aim to reach the consumer, which causes a great increase of integrating their brand and products into non-traditional formats (Boerman, van Reijmersdal & Neijens, 2012). One of these formats is the use of YouTube videos, where a YouTuber is contacted by a brand to collaborate with them on a project. The YouTuber will then make a video that will include or highlight the product or brand in some way, reaching their followers, whilst he/she is being paid to make this video. This type of brand integration is often seen as the future of advertising (Hudson & Hudson, 2006). There has been a lot of debate about the ethical issues that come with sponsored content. The persuasive nature of this type of communication is often considered deceptive, since it is harder for consumers to distinguish between commercial and noncommercial content (Nebenzhal & Jaffe, 1998). This may consequently result in less critical processing of the message, which in turn 'could lead to unaware persuasion' (Boerman, van Reijmersdal & Neijens, 2012, p. 1048). There is a 'simple' way in which these ethical issues can be solved. By telling their viewers that

they are or are not being sponsored by or have a coupon code for a specific brand, YouTubers are disclosing the fact that they are being paid and/or are receiving commission for products that are sold through them. However, as mentioned before, this disclosure may also bring some problems, as it has proven to have a negative effect on brand attitude (Wei, Fischer & Main, 2008). This study attempts to find out what exactly the effect of this disclosure is on the effectiveness of the advertisement that is featured.

2. Theoretical framework

In this chapter, the various theoretical concepts that this research involves will be discussed. Together, the discussion of these concepts will provide a clear expectation of the answer to the research question. First of all, the basic concepts that surround this research will be discussed. This includes user-generated content, online marketing and YouTube as a platform. These sections will map out the direction of the research more clearly. Afterwards, a narrower perspective is taken as the beauty industry is highlighted and discussed. In the next section, the main topic of the research, native advertising on YouTube, is discussed. This section also includes the formulation of the hypotheses. The sixth and final section discussed the broad operationalization of advertising effectiveness, the variable that will be used to be able to accept or reject the hypotheses.

2.1 The Rise of the Web 2.0 & User-generated Content

Social media has left and is still leaving an inerasable impact on the world we live in. It is striking to think about how social networking sites like Facebook and YouTube were only established some 13 years ago (Ellison, 2007), but to many people, especially the younger generations, it is like they have always been around. The internet as we know it today is mostly associated with the concept of Web 2.0, a term popularized by Tim O'Reilly in 2004. However, the phrase was first used in 1999 by information architecture consultant Darcy DiNucci, who wrote about how the Web was only in its first stages and would soon progress to a much more interactive 'transport mechanism' (DiNucci, 1999, p. 32). And indeed, soon the web would evolve to a network-based platform 'spanning all connected devices' (O'Reilly, 2007, p. 17). Within O'Reilly's view of the new Web there was a central position for the individual user, which provides its 'own data and services in a form that allows remixing by other' (p. 17). This is also called user-generated content (UGC). UGC began with the rise of web logs ('blogs') in 1999 (Williams & Jacobs, 2004). People 'voluntarily contribute data, information or media' (Krumm, Davies & Narayanaswami, 2008) to online platforms with the goal of informing or entertaining. Moreover, collective intelligence was another key point in O'Reilly's manifest, which he explained by example of websites like Flickr and Wikipedia. Users can add, remove and mix data and content and in this way build a collective network of knowledge and information through these platforms. Altogether, the Web 2.0 was responsible for the headstart of user-generated content.

2.2 Online Marketing & Advertising

The rise of the Web 2.0 did of course not go unnoticed by businesses. During the dotcom bubble, commercial interest in the Internet began, and it has been there ever since. The emergence of social media was a great opportunity for companies wanting to advertise their product to a global, yet segmented audience. Even though social media are often applauded and put on a pedestal for 'facilitating political and cultural participation' (Crain, 2014, p. 371), these platforms are created and exploited by companies seeking to monetize the every move of users. YouTube, Twitter and Facebook live by the grace of companies that want to advertise their products on the platforms. There are three main things that separate the internet from other communication channels that are used to spread advertisements (Gurău, 2008). The first is interactivity, which comes down to the fact that the internet acts both as communication agent and as interface or platform for other things to happen on. The second is transparency, which relates to the ability of any Internet user to access information that was published online, unless of course it is specifically protected. The last characteristic that seperates the Internet from other communication channels is its memory, as the Internet both transmits and stores information at the same time. The thing that marketers can profit from the most with regards to online marketing, is the fact that user's behaviors are completely trackable and can be analyzed with the push of a button. The so-called Big Data streams, which are vast amounts of information that can be analyzed instantly, can help marketers that seek to find out more about the behavior of Internet users; when do they click on an advertisement? How long do they spend on a certain page? What types of colors, fonts, and placements of titles trigger the best reaction? These example questions all show how much information there is to process, and how much companies can increase their advertising effectiveness if they 'listen' to these data streams (Chen, Chiang & Storey, 2012)

It is logical that as social media and the internet further develop further, advertisement forms develop with it. A new form of marketing within the online sphere is influencer marketing. Influencers are people who have an extensive amount of followers on social media platforms (mainly Instagram and YouTube) (Perloff, 2014). Some influencers are celebrities in other industries (movies or music), and some influencers are merely well-known through social media and this is also how they make their living. The term 'influential' was already used a long time ago to describe people with high leverage within their community, but with the rise of social

media the extent of this community have expanded far beyond personal contacts. A lot of these influencers have a couple of million followers, and people are able to follow their every move through Instagram posts, live Snapchats and daily video diaries on YouTube. Their followers often see these influencers as examples, and want to have what they have. Therefore, companies have started to use influencers as advertising machines: the influencers get paid to show the product, talk about it and/or recommend it to their followers (Goulston & Ullmen, 2013). The downside and ethical issue of influencer marketing is that followers are persuaded to buy something because they trust the person that uses it, even though they are being paid to feature the product on their social media accounts (Wright, 2017). Therefore, marketers and influencers alike have to start wondering when and how is the best and most consumer-friendly way to advertise the products.

2.3 YouTube & Advertising on YouTube

YouTube is one of the pioneers of the Web 2.0. It has been the biggest online video-sharing platform since its establishment in 2005 (Ebizmba.com, 2017). The website has gone from 100 million views per day in 2006 (Gill, Arlitt, Li & Mahanti, 2007) to 5 billion in 2018 (Digital Marketing Education, 2018). YouTube is what is called a 'content community' (Kaplan & Haenlein, 2010), which has a main objective of sharing media content between users. Next to news and professional music videos, most videos on YouTube are categorized as user-generated content; people record a video themselves, edit it, and put it online through their YouTube channels. The influencer marketing technique is very much exploited on YouTuber, as well as regular advertisements before and during videos. Product placements, mainly in vlogs, is another way in which marketers aim to advertise their product to target audiences (Gerhards, 2017). In general, there has not been an extensive amount of research about YouTube, let alone research about YouTube in the context of marketing and advertising. This may also be due to the fact that companies have not yet fully realized the potential of the platform (Reino & Hay, 2016).

A theory that explains the popularity of influencers on YouTube very well is the social cognitive theory, developed by Bandura (1986). The theory holds that the observation of others ('models') is a key influence in the performance of individual behaviors. In this case, YouTubers can be considered the models for the consumers. People identify themselves with these YouTubers and adopt their values, attitudes, beliefs, and behaviors. This identification is often based on interests, lifestyles, and gender, which is similar to audiences that watch beauty content

on YouTube. Since there is so much content of these YouTubers available for the consumer to see, it becomes much easier to influence them and therefore the effects of exposure are likely bigger than for example for celebrities that people only sometimes see in commercials or a movie.

The concept of influencer marketing is certainly very applicable to YouTube marketing techniques. Influencer marketing is a type of 'virtual word-of-mouth' marketing (Woods, 2016; p. 6), which has a much greater effect on consumers than most other types of marketing. This is because the message comes from a person that the audience 'trusts', that is; the person is a friend or someone that they follow on social media, so-called 'influencers'. According to a Twitter study, 56% of social media users rely on recommendations from friends if they buy a product, and 49% say they rely on influencers (Swant, 2016). Companies are always looking to build better relationships with consumers, and if they cannot do so directly they are more than willing to do it through influencers. If your favorite influencer shows the same type of cereal everyday in his vlogs, you are also going to be more likely to buy it. You trust him and if he likes the cereal, it's likely that you will too. With influencer marketing, the most important thing is to identify the decision makers within the target audience. For beauty companies, the target audience gets their information mostly on Instagram and YouTuber, and therefore using these beauty gurus and reaching the audience through these platforms is the appropriate technique (Jaakonmäki, Müller & vom Brocke, 2017).

2.4 The Beauty and Cosmetics Industry

The global cosmetic market grows with about 4% every year. The Asian-Pacific market is the most profitable, owning about 38% of the total worldwide market in 2016 (L'Oréal, 2017). L'Oréal is by far the biggest beauty manufacturer, with a yearly revenue of 28.6 billion dollars (Women's Wear Daily, 2017). The U.S. cosmetic market is the most profitable, with a value of over 80 billion dollars in 2015 (Euromonitor, 2016). Within the beauty industry, the biggest market share is for skincare products (about 36.4%), after which come haircare (22.9%) and makeup (18.2%) (L'Oréal, 2017). The size of the make-up market in 2016 was 64.1 billion dollars worldwide (Trefis, 2017). In 2015, 51% of the beauty content videos were about makeup products, the #1 beauty product category on the platform (Pixability, 2015). The most mentioned beauty brands mentioned by vloggers in 2015 were MAC (9.6% of the mentions), NYX (3.7%) and NARS (3.2%). Beauty creators made almost all of the content about beauty on YouTube in

2016 (97.4%) compared to the beauty brands themselves (Pixability, 2016). This already proves the importance of beauty influencers within the cosmetic industry. Beauty influencers control the market on YouTube! In 2015, 45% of beauty content videos were labelled as 'tutorial', 20% as 'commercial', and 5% as 'product demo' (Pixability, 2015). However, it seems as if the meaning behind these titles is not quite clear: is a video where someone creates a look using products that they are reviewing at the same time a product demo or can it also be called a tutorial? And what is the difference between a tutorial, a routine, and a look? The discrepancy between these terms may have an effect on how a consumer looks at the video and why he/she decides to click on it. However, this research topic has not been extensively researched yet.

      With regards to consumer behavior, in a representative sample of U.S. consumers, researchers found that 35.5% of US consumers state that they are influenced by positive or negative reviews for their decision to buy or not buy cosmetic products. Moreover, 27.7% of consumers thoroughly research products before they try them. Another 18.8% of consumers often tries products that they see demonstrated and/or review on social media (Toluna Quicksurveys, 2016). Make-up can thus be seen as a medium to high-involvement product: a lot of people do not just go out and buy a random product, but they make conscious and evaluated decisions. Coupon codes are an easy and often-used way for social media beauty influencers to market a product towards their followers. These influencers get an amount of money based on how many people use the specific coupon code. 21% of U.S. consumers frequently uses these online codes, and 37% uses them occasionally (Statista, 2017). When asked how often they watch product reviews and tutorial videos respectively, 22% and 19% of consumer stated they regularly watch them. 32% and 34% stated they watch them occasionally (Statista, 2017). One of the most common types of videos in the beauty and fashion communities on YouTube are haul videos. In these videos, the YouTuber shows and talks about new items that they recently purchased, often resulting in them unpacking entire boxes and big bags of products. 62% of all consumers stating they have seen these haul videos once or more. Moreover, 40% consumers say that they prefer beauty videos personal accounts over those by cosmetics companies, against 29% preferring cosmetics companies (Statista, 2017). However, there is much less content available from the cosmetics companies, as stated before. All in all, it is easy to see just how big the cosmetic industry is, and how important its stakes in the social media sphere are. YouTube is

one of the most important platforms for beauty companies to feature their product on, and so there is a high need for more research on this topic.

2.5. Native Advertising on YouTube

2.5.1. Sponsorships & coupon codes

YouTube has proven to be an excellent platform for companies, those the beauty industry in particular, to promote their products. YouTubers put videos online in which they either perform an activity whilst being sponsored by a company, for example for a company like GoPro, or they show how a product works in the video whilst being sponsored, which is often the case in the beauty industry. Next to that, YouTubers receive free PR packages with goods from companies, and the companies hope that they will use or unbox them on screen. Lastly, YouTubers make a deal with companies that they will spread a coupon code for a product from that company for the viewers, and they get paid a share of the price of the products through commission. The more people use the coupon code, the more money the YouTuber gets. This type of marketing fits perfectly under the description of native advertising, as it is a form of paid advertisement presented as editorial content by a publisher (Wojdynski & Evans, 2016), in this case the YouTuber.

Native advertising has become an important issue in the age of consumer-generated content, since the boundaries between advertising and personal opinion, especially in videos where a product is 'reviewed' have become blurred (Wu, 2016). The disclosure of sponsorships on YouTube has always been a difficult topic, since it does not apply strict guidelines for sponsorships on YouTube; YouTube does not require creators to disclose sponsorships. This makes financial interest often go well above viewers' trust, as YouTubers positively review products because they get paid for it. This is especially concerning when it comes to the youngest generations of YouTube users, as they are not yet fully aware of advertising intentions and may have a hard time spotting the difference between paid and non-paid content if this is not disclosed. It is a known fact that generally, consumers react negatively when it is disclosed that something is an advertisement (Wojdynski & Evans, 2016). Moreover, the credibility of the source and/or the message of content is evaluated as less credible once the use of native advertising is disclosed (Carlson, 2015). This together makes it very hard for YouTubers who want to earn money with their videos to balance money and trust.

Boerman, van Reijmersdal and Neijens (2014) investigated if the disclosure of a sponsorship influenced the persuasive effect of the content featured. The results of this study showed that disclosure of sponsorships negatively influenced the brand attitude of the brand featured. They proved that disclosure can make people aware of sponsored content and thus review it more critically. More importantly, it can lead to a resistance towards the persuasion and thereby negatively impact advertising effectiveness. Aqsa & Kartini (2015) also recognized the importance of disclosure of advertisements. Overall, literature on disclosure of sponsorships and advertisements is quite scarce (Hoofnagle & Meleshinsky, 2015), mainly because it has only become an important way for marketers to persuade consumers over the past decade.

Based on expectations from literature, the following hypotheses were formed. All hypotheses that talk about disclosure of (non-)sponsored videos and/or coupon codes are about visual disclosures in the description boxes.

H1a: The disclosure that the video is sponsored will result in less advertising effectiveness than the disclosure that the video is not sponsored.

H1b: The disclosure that there is a coupon code for the product featured in the video will result in less advertising effectiveness than the disclosure that the video is sponsored.

H1c: The disclosure of a sponsorship or a coupon code will result in less advertising effectiveness than an empty description box.

H1d: An empty description box will result in less advertising effectiveness than the disclosure that the video is not sponsored.

2.5.2. Tutorial vs. review

Within the beauty community on YouTube, there are several types of categories that videos can belong to. There are challenges, routines, unboxings, and many others. The two most often used titles for beauty videos have to be tutorials and reviews. Tutorials are videos in which the YouTuber explains how he/she uses one or multiple make-up products. For example, a video could be titled: 'Tutorial: How to Apply a Cat-Eye Eyeliner Perfectly'. Reviews are videos in which the YouTuber explains what he/she thinks of a new product, or at least a product that they have not used yet. The main difference in terms of marketing possibilities with these videos is that reviews are usually supposed to be unbiased and just a clear and honest opinion about a product. When YouTubers would be paid to give a review, it would immediately not be as unbiased anymore. However, in the beauty community, these two types of titles are used

11

interchangeably for videos that have and do not have sponsorships within them, whether it's disclosed by the YouTuber or not.

Based on the previous statements, the following hypotheses were formed.

H2a: Overall, a video with the title of 'tutorial' will result in less advertising effectiveness than a video with the title of 'review'.

H2b: A video with the title of 'review' will only result in less advertising effectiveness than a video with the title of 'tutorial' if the description box either states that the video is sponsored or contains a coupon code.

2.6. Advertising Effectiveness

The end goal of every advertising or marketing effort is to get more people to buy a product or use a service. The construct through which this goal is operated is called advertising effectiveness. In order to understand the purpose of this research, it is important to look into the concept of advertising effectiveness and disassemble it in order to fully understand its scope. It is important to point out the discrepancies of the term advertising effectiveness, since there are several theories and models that explain the term with various variables. Native advertising is a type of advertising, of course, but since the research also concerns a condition in which it is apparent that there is not a single trace of advertising (conditions with 'this video is not sponsored'), the term 'advertising' may not be fully accurate. However, for the sake of the majority of the conditions, the advertising effectiveness will be used as a general term to describe the effects of the videos on the consumer's ideas and perceptions about the brand featured.

For the sake of this research, advertising effectiveness will consist of two main variables: brand attitude change and purchase intention change. The three main variables that will likely be able to explain certain outcomes of brand attitude change and/or purchase intention change are source credibility, message credibility and video liking. The construction of advertising effectiveness is based on a model developed by Ko, Kim, Taylor, Kim and Kang (2007). In their research, lifestyle and nationality predicted the variables of advertising effectiveness. In this particular research, the context of the YouTube video will predict how effective the advertising is. The elements comprising advertising effectiveness will be discussed next.

2.6.1. Brand attitude

According to Fishbein and Azjen (1975, p. 222), "a person's attitude is a function of his salient beliefs at a given point in time.". There are several models that attempt to explain brand

attitude. A very common theory is that of the Elaboration Likelihood model by Petty & Cacioppo (1986). The Elaboration Likelihood Model (ELM) is a predictor of attitude change. The two most influential factors in this model are motivation and ability to process the message. The authors state that there are two ways in which a consumers can have an attitude change; they can take either the central route or the peripheral route. The central route is one where attitude change comes from a person's logical reasoning of why they have a particular attitude (usually of a product/service). This route is mainly based on cognitive theories, and attitude changes that occur through this route are usually enduring and 'predictive of behaviour' (Petty, Cacioppo & Schumann, 1983). With the peripheral route, on the other hand, attitude changes occur because the issue or object are linked to positive or negative cues in the context of the persuasion. Their study showed that the peripheral route of attitude change is often taken. It will be very interesting to see how the attitude a person has towards the brand can affect the processing of the video itself, and the effectiveness of the video itself as well.

2.6.2. Purchase intention

Purchase intention measures the possibility that a consumer will buy a product (Schiffman & Kanuk, 2000). Purchase intention is a more difficult to interpret, since its role varies in different theories and models. For all advertising and marketing efforts, the goal that is closest to actual purchase of the product, the action which itself is not controllable anymore by marketers, is purchase intention (Aqsa & Kartini, 2015). There are several ways in which purchase intention can be influenced, either directly or through other variables. Sometimes, purchase intention is seen as the end step in a sequence of brand familiarity and brand attitude. However, in this research it is seen as a variable that is parallel to brand attitude, rather than it being an effect of it. It is also being measured separately.

2.6.3. Source credibility

The variable of credibility will consist of the two sub-variables of source credibility and message credibility. Source credibility has long been proven to be a major factor in establishing communication effectiveness (Hovland & Weiss, 1951), and soon after the start of advertising research in establishing advertising effectiveness. This is one of the reasons why certain types of people are employed for certain types of advertisements; household product often have people that are similar to the target audience and luxury and beauty products often have celebrities. These people seem more credible towards the target audience. The concept of source credibility

consists of two dimensions: (1) expertise, 'the degree to which the source's knowledge is offered as valid information', and (2) trustworthiness, 'the honesty and morality of a source in providing objective and unbiased information' (Wu et al., 2016). Highly credible sources elicit more desirable attitude changes than less credible sources. Moreover, source credibility has also proven itself to be an influence on purchase intention (Lafferty, Goldsmith & Newell, 2002). Therefore, it can be expected that more source credibility will result in more advertising effectiveness. Even though scholars often consider source credibility to be one variable, it has been proven that the breakdown of source credibility into source trustworthiness and source expertise is important (Lemanski & Hyung-Seok, 2012), as their effects on attitude can be quite different.

2.6.4. Message credibility

Message credibility is based on the variable 'ad credibility', developed by Lutz (1985, as cited by Kim & Damhorst, 1999). Ad credibility was defined as 'the extent to which the audience perceives claims made about the brand in the ad be truthful and believable' (p. 20). Ad credibility is conceptualized to be a main determinant of ad attitude. For this research, both terms were slightly changed to fit the research design; since we are talking about a video, and not an actual advertisement, message instead of ad credibility seemed like a better option. The construct of the variable does not differ from that of ad credibility, however (Lutz, MacKenzie and Belch, 1983).

2.6.5. Video liking

The variable video liking is based on the variable 'ad liking', and its name was slightly adapted for the previously mentioned reason. Researchers have also studied ad liking (or 'advertisement likeability') as a variable that affects the effectiveness of sales efforts (Tomkovick, Yelkur, Christians, 2001). A research by Biel and Bridgwater (1990, p. 40) showed that 'people who like a commercial are twice as likely to be persuaded by it than people who simply felt neutral towards the advertisement'. Moreover, Homer (1990) stated that when marketers expect consumers to go through a logical decision-making process when they are evaluating a brand, designing a well-liked advertisement is even more important. An ad that is well-liked is proven to perform better in terms of perception and persuasion (Stapel, 1994). Ad liking can be compared to attitude towards the ad, a common mediator of advertising effectiveness (Lutz, MacKenzie & Belch, 1983)

3. Method

This study aimed to find out what effect the context of YouTube videos has on the effectiveness of advertising. In order to research this, a quantitative method using an online survey and experiment was put out towards the sample population. Quantitative methods aim to explain phenomena, in this case what factors influence advertising effectiveness, and they are based on collecting numerical data which is analyzed mathematically (Muijs, 2010). The results of this research, based on the answers given in the survey, were numerical and therefore open for an analysis through a statistical program in this case with SPSS.

3.1 Research Design

In order to answer the research question, this research was conducted through a 2x4 factorial quasi-experiment, based on an online survey. The research tested whether the context of a YouTube video (the title and content of the description box) had an effect on the effectiveness of the advertising within the video. The main objective of the study thus was to see whether respondents show a statistically significant change in their answers to questions about credibility of the source, brand attitude and brand recall. Below are the identified conditions that the respondents could have been exposed to.

As mentioned, the research was an online experiment, in the form of a survey. The choice of an experiment was based on the fact that the research aims to measure advertising effectiveness, which means that there is a need for both the determination of causal relationships (e.g. between the title of a video and brand recall), but also for a research environment that is not as controlled as a laboratory experimental setting (Goldberg, 1990). The online survey as a response mechanism was therefore the most ideal, since it can be filled out with ease and is easy to spread (Wright, 2005). Moreover, it gives access to the desired population within an instance, since YouTube channels can function as community gathering places for the like-minded.

It is important to also mention the disadvantages of conducting an online experiment. First and most important of all, the most prominent disadvantage of online experiments is the fact that they are extremely difficult to control. Respondents can conduct the experiment where and whenever they like. This can increase variability, resulting in a possibly less reliable outcome (McGraw, Tew & Williams, 2000). A big enough sample size can compensate for this, but it will always remain a factor to keep in mind. Another disadvantage of experiments, online or offline, is the ethical concerns that come with it. Even though the researcher will always

explain the circumstances of the experiment and needs an informed consent from the respondent before the experiment can begin, ethics remain a difficult part of this type of quantitative research. What can and what can't a researcher do or ask from their respondents? Another disadvantage with online experiments specifically is that respondents have the ability to fill out, in this case, the survey, resulting in less reliable results. Since this online experiment was strictly confidential, the only way in which this can be prevented was by filtering out responses with the same IP-address (Bryant, 2004). For this particular research, there was a problem with the drop-out rate, another common disadvantage of online experiments in particular. Just over a third of people who started the survey fully finished it. However, the researcher was able to continuously monitor the response rate so that enough respondents could still be gathered.

3.1.1 Stimulus material

The stimulus material in this research was a single video that was created especially for this study. The researcher contacted the YouTuber if she would like to collaborate for the research, and they sat down together to discuss what elements should be in the video. A script was written for the YouTuber to follow, and subsequently the video was recorded in collaboration with the researcher. The fact that the video was made solely for the purposes of this study made it possible to control all elements of the video and is quite unique. For the survey, the same video was uploaded on YouTube 8 times with a different title and description box. The video was recorded by YouTuber Britt Alsemgeest, known on YouTube as Basically Britt. She uploads about one or two times a week, and all of her videos are recorded in English, even though she is Dutch. Britt has nearly 6,000 followers, which made the video a lot more believable for the respondents than if the video were to be uploaded through a 'new' account with 0 followers. The difficulty with a video in a survey is that respondents are often distracted or do not like to watch the whole video, especially not if it is too long. However, beauty videos are often as long as 8-10 minutes. It was therefore crucial to keep an aspect of realism to the video, without boring the respondent. Therefore, all that respondents got to see is a recorded screenshot of a video on YouTube, where it appeared that the video was 8 minutes long, but they only got to see an extract of this video of 1.5 minutes. In reality, if the video were to have played on it would just be a black screen with no sound. In this way, boredom or distraction was avoided as much as possible, without compromising the realistic nature of the video. The product featured in the video was the Dream Satin Liquid foundation by Maybelline. It was chosen to use

a foundation for the video as it often takes a while for someone to put it on their face, and it was important that the video was not too short. If the product were to have been a lipstick or mascara, the video would have been a lot shorter. Moreover, foundation is one of the most basic and common make-up products for people to start their make-up routine with, and thus it was most likely that the viewer could identify themselves as they would also use foundation themselves. Before the video started, a small text explained that the respondents were about to see an extract of a YouTube video. Since this research allowed respondents of any nationality, it was important to use a well-known international brand that does not cost too much but does have a good reputation in the beauty community. Maybelline has all of these characteristics. The disadvantage of using Maybelline was that people likely had an opinion about the brand and possibly the specific product beforehand, which could have influenced results. However, if an unknown brand was chosen, there could be a risk of low purchase intention scores in general, as low brand familiarity causes purchase intention to also be low.

Since the video had to be able to represent both a tutorial and a review, as it can have both of these titles, it could not lean too much towards either of these types of videos. A tutorial would usually be a video in which the YouTuber explains how he/she does something, for example apply eyeliner or make your lips look fuller. A review, on the other hand, would be a video where the YouTuber talks about a new or unused product, and in which he/she gives a final opinion on the product. These videos may seem very different, but in reality it turns out that YouTubers use these terms interchangeably. However, in order to minimize the potential impact of the video itself and rather have the title and description box be the actual stimulus of the experiment, the following 'script' was developed and made into a video:

1. The video is a screenshot of a YouTube video's page. The respondents could see everything that a 'normal' audience would also be able to see; the title, the amount of likes/dislikes, the amount of subscriber the YouTuber has, the YouTuber's name, the amount of views, and the content of the description box. When the video started playing, the time display in the YouTube video showed that the 'actual' video lasts for about 8,5 minutes. However, respondents of this survey only got to see 1,5 minutes of this. This embedding was necessary because the videoplayer in Qualtrics does not allow to show the YouTube framework around it, and linking the respondents to a YouTube page would

only distract and could possibly lead to them clicking another video or entirely quitting the survey.

2. Britt welcomed the viewers, and told her viewers what she would be doing in the video. She mentioned the name of the product ('Maybelline Dream Satin Liquid Foundation') and got to the point fairly quickly, mostly to avoid respondents to be bored fast.

3. Britt talked a bit about her skin type and what other products she used before, which is a common thing for beauty YouTubers to do.

4. Britt went through all the steps in which she applies the foundation, and during these steps she also described how the foundation looked, felt, applied, etc.

5. Britt was quite positive about the product, and discussed why. She mentioned the color name and code of the foundation, another common feature in beauty videos.

6. Britt further evaluated the product and mentioned what type of people could also benefit from the product.

7. Britt mentioned how she was now done with her foundation and would move into another step of her make-up routine process. The respondent was made to believe that Britt would continue, but the video stopped. This was the end of what the respondents get to see.

3.2. Sampling Method

For this research, the sampling frame consisted of females from any country between the ages of 16 and 55 who had whatever type of interest (a lot or a little) in beauty and/or make-up. All responses were gathered through the use of social media and platforms like YouTube, Reddit and Facebook. Therefore, all respondents had to have an account on either one of these platforms in order to be able to participate. Since women are typically more interested in the beauty products on display, the choice was made to limit the sample to women. Since only a few men (n=9) participated in the survey, this decision helped to minimize noise. Even more so, marketers in the beauty industry are also still focusing on women as their main target group. With this, we also avoided that men would just fill out the survey because they were helping the researcher, instead of filling the survey in in a truthful manner. Moreover, respondents younger than 16 and older than 55 were excluded. Respondents younger than 16 are typically not the main audience for videos like the one featured in this video, and are therefore not an important audience for marketers in the beauty industry that would want to know more about audience's behavior. Respondents older than 55 were also not included in this research, since they fall outside the

target audience of this product. Again, this research aimed to provide more insights for beauty marketers, which is why the sampling frame is narrowed down to specific ages that correspond with the marketers' interests and target groups.

This aimed sampling size was determined by multiplying the number of conditions, eight in this research, by the minimum amount of respondents for experiments in a Master Thesis according to the Methodological Guidelines Thesis Research (ESHCC, 2017) which is 30 respondents. Of course, having more respondents than required only added to the reliability and representativeness of the research.

The sampling method used in this research was not random, rather it was done through homogenous purposive sampling. Purposive sampling is a non-probability method based on the notion that respondents are selected deliberately, due to the quality that the participant possesses (Etikan, Musa & Alkassim, 2016). In this case, these qualities were gender, age and an interest in make-up and beauty. This was a necessary precaution, since factors like disinterest in the subject had be controlled for in order to stop them from intervening with the results. Finding people that are genuinely interested in make-up greatly increases the validity of the research. The sampling method is called 'homogenous', because the research looks at people with similar traits. All of the respondents were collected through various media on the Internet, as this is recognized as the fastest and easiest way to reach people with the characteristics that were required for this study (Wright, 2005).

3.3. Sample

In total, 707 people started the survey. Of course, this data had to be cleaned. As per usual, not every respondents fully finished the survey. For most this was because they simply quit their participation midway through the survey. For others this was because they were not able to see the video (n = 88), in which case the respondents were directed straight to the end of the survey. Looking at the reasons why these people were not able to see the video, which they had to indicate, this was mostly due to technical issues. There were 289 people who managed to see the video and make it to the last question. However, this sample size included people who were either too old, too young, or of the male gender. These respondents were removed from the sample. The valid sample, therefore, consisted of 258 respondents (N=258). These respondents came from 19 different countries, including 13 European countries. Most respondents came from the Netherlands (75.2%), followed by the United States (10.5%) and Belgium (3.1%). The

division between Dutch respondents and those of other nationalities is likely due to the fact that the request for participation were posted under a lot of Dutch YouTube videos, together with the fac thtat the researcher was Dutch and posted the survey on her Facebook page. Respondents were between the ages of 16 and 54 years old, and they had a mean age of 23.13 ($SD = 7.29$). The distribution of age among respondents was not quite equal, as 47.8% of respondents (n=110) were between the ages of 16 and 19 years old, 40.4% (n=93) were between the ages of 21 and 25 years old, and only 11.7% (n=27) of respondents were 26 years old or older. This was likely due to the fact that the survey was spread on YouTube as its main medium, which is a relatively 'new' medium that is mostly used by younger age groups.

As mentioned before, each of the respondents was shown one out of eight different stimulus videos. Below a table is given that illustrates the division of respondents between each condition, as well as the eight possible conditions the respondents could be exposed to.

Table 3.1: Representation of (respondents within) each of the eight conditions

	Review	Tutorial	Total
Coupon	12.4% (32)	13.6% (35)	26.0% (67)
Empty	11.6% (30)	14.7% (38)	26.3% (68)
Not Sponsored	11.2% (29)	12.4% (32)	23.6% (61)
Sponsored	12.8% (33)	11.2% (29)	24.0% (62)
Total	48.0% (124)	51.9% (134)	100.0% (258)

3.4. Operationalization & Measurements

As mentioned, the end goal of this research was to find out how advertising effectiveness is influenced. The concept of advertising effectiveness is very broad and can be explained through a number of dependent variables. For the sake of this research, it was decided that the two variables that would measure advertising effectiveness were going to be brand attitude and purchase intention. In order to explain the variance in these variables, another three variables were used. These variables include source credibility, message credibility and video liking.

3.4.1. Brand attitude

In the research, brand attitude towards the product brand that was featured in the video, namely Maybelline, was measured according to the seven-point scale with five semantic differential items, based on the article by Spears and Singh (2004). The original scale has often been proved to be highly reliable (Cronbach's α = .97 (Spears & Singh, 2004)). Each respondent got to see a question with five items before and after they have seen the video, and they were asked to fill each item on a scale, on one of the seven points, for example ranging from bad to good or from unlikable to likable. There was no 'true' control group in this research, as this was difficult to accomplish since YouTube videos work with a standard format. The video is required to have a title, and so the conditions with the empty description box basically acted as control groups. Since the research aimed to study the change occurring in respondents' brand attitude and purchase intention because of the video that they saw, respondents had to answer this question about brand attitude and purchase intention both before and after the stimulus material. For the measurements of brand attitude, a new variable – *Brand Attitude Change* – could be created for analysis by subtracting the scores of the pre-video brand attitude questions from the post-video brand attitude questions.

For both pre- and post-video brand attitude, principal component analysis showed that all five items of the brand attitude scale loaded on a single component (pre-vid: Eigenvalue 4.08, post-vid: Eigenvalue 4.21). The Kaiser-Meyer-Olkin value of .90 ($p < .001$) and .90 ($p < .001$) respectively supported the combination of these items. Together, the items explained 81.5% of the variance in pre-vid brand attitude (factor loadings as follows: .87 for Unappealing-Appealing, .90 for Bad-Good, .94 for Unpleasant-Pleasant, .90 for Unfavorable-Favorable, .90 for Unlikable-Likable). For post-vid brand attitude, the items explained 84.2% of the variance in post-vid brand attitude (factor loadings as follows: .91 for Unappealing-Appealing, .92 for Bad-Good, .94 for Unpleasant-Pleasant, .90 for Unfavorable-Favorable, .92 for Unlikable-Likable). A reliability analysis was conducted to confirm internal consistency for both pre-vid brand attitude (Cronbach's α = .94) and post-vid brand attitude (Cronbach's α = .95). Finally, the overall score for the *Brand Attitude Change* variable was calculated by averaging scores for items of pre- and post-vid brand attitude, and subtracting the average pre-vid brand attitude score from the average post-vid brand attitude score.

3.4.2. Purchase intention

Purchase intention was also measured according to a scale developed by Spears and Singh (2004). It concerned a seven-point scale which had to be answered based on the question 'How likely are you to buy a primer and/or foundation from Maybelline?'. The question for purchase intention was, like with brand attitude, asked twice; before and after the stimulus video. This was because the research aimed to measure the change in purchase intention that was caused by the stimulus material. The new variable of *Purchase Intention Change* was created by subtracting the pre-vid purchase intention scores from the post-vid purchase intention scores.

3.4.3. Determining variables for advertising effectiveness

Source credibility was measured through the use of two items, based on Ohanian's understanding of source credibility (1990). The first variable to measure source credibility was *Source Trustworthiness*, and was measured by having respondents stating how much they agreed with each of the five statements based on a five-point Likert scale (Likert, 1932). An example of such a statement is 'The Youtuber in this video is ….'. The five items measuring source trustworthiness were as follows: (1) honest, (2) dependable), (3) reliable, (4) sincere, (5) trustworthy. The second variable to be measured was S*ource Expertise*. This scale was proven to be highly reliable (Cronbach's α = .94). Again, there were five statements which were based on Ohanian (1990) and were to be answered through a five-point Likert scale. An example of a source expertise statement was 'The YouTuber in this video is ....'. The five items measuring source expertise were as follows: (1) knowledgeable, (2) skillful, (3) experienced, (4) qualified, (5) an expert. This scale has also proven to be reliable (Cronbach's α = .90).

Principal component analysis showed that all five items on the scale of source trustworthiness were loaded on one component (Eigenvalue 3.55). The Kaiser-Meyer-Olkin value of .87 ($p < .001$) supported the combination of these items. Together, the items explained 71.1% of the variance in source trustworthiness (factor loadings as follows: .87 for Honest, .73 for Dependable, .89 for Reliable, .85 for Sincere, .87 for Trustworthy). A reliability test was conducted to confirm internal consistency (Cronbach's α = .90). Finally, the overall score for source trustworthiness was created by averaging scores for all five of the sub-scale items.

Principal component analysis showed that all five items on the scale of source expertise were loaded on one component (Eigenvalue 3.16). The Kaiser-Meyer-Olkin value of .81 ($p < .001$) supported the combination of these items. Together, the items explained 63.2% of the

variance in source expertise (factor loadings as follows: .75 for Knowledgeable, .83 for Skillful, .85 for Experienced, .81 for Qualified, .73 for An Expert). A reliability test was conducted to confirm internal consistency (Cronbach's α = .85). Finally, the overall score for source expertise was created by averaging scores for all five of the sub-scale items.

The credibility of the message was measured using a newly-developed scale by Appelman and Sundar (2016). This scale for message credibility consists of three adjectives as items: accurate, authentic and believable. The scale has proven to be reliable in the past (Cronbach's α = .95). The respondents were asked to indicate how well the three adjectives described the content of the video. They were able to answer on a five-point scale ranging from 1 = *extremely well* to 5 = *not well at all*. It was decided to also use an additional seven-point bipolar adjective scale with four items: (1) biased/unbiased, (2) deceptive/not deceptive, (3) honest/dishonest and (4) convincing/unconvincing (Kim & Damhorst, 1999). This scale, too, has proven itself to be highly reliable (Cronbach's α = .93 (Kim & Damhorst, 1999). Not only did the use of two scales instead of one scale improved validity of the experiment, but it also showcased the value of the scale by Appelman and Sundar. Since the first two items of the Kim and Damhorst (1999) were based on a scale from negative to positive (biased to unbiased, deceptive to not deceptive), and the last two items were based on a scale from positive to negative (honest to dishonest, convincing to not convincing), it was decided that item 1 and item 2 had to be recoded so that these variables would also have a scale from positive to negative. This was also decided due to the fact that they would likely combine with the scale of Appelman and Sundar (2016), whose items also were based on a scale from positive to negative.

When combining all of the seven the items of both scales, principal component analysis showed that all of the items of the Appelman and Sundar (2016) together with item 3 and item 4 of the Kim and Damhorst (1999) scale loaded on factor (Eigenvalue 2.93), and item 1 and 2 of Kim and Damhorst (1999) loaded on another factor (Eigenvalue 1.35). The Kaiser-Meyer-Olkin value of .73 (p < .001) supported the combination of these items. Component 1 explained 41.8% of the variance in message credibility (factor loadings as follows: .69 for Accurate, .63 for Authentic, .70 for Believable, .76 for Honest/Dishonest, .82 for Convincing/Not Convincing). Component 2 explained 19.3% of the variance in message credibility (factor loadings as follows: .86 for Unbiased/Biased, .86 for Not Deceptive/Deceptive).

The reliability analysis for the first component of message credibility confirmed internal consistency for this variable (Cronbach's α = .74). The overall score for this variable was calculated by adding up the five items and dividing the sum by five. Since all of the items that this variable consists of are positive connotations of message credibility, this variable was called *Positive Message Credibility.* The reliability analysis for the second component of message credibility confirmed internal consistency for this variable (Cronbach's α = .74). The overall score for this variable was calculated by adding up the two items and dividing the sum by two. Since both of the items that this variable consists of are negative connotations of message credibility, this variable was called *Negative Message Credibility*.

The final determining variable that was used to explain possible effects on advertising effectiveness was video liking. Video liking was measured based on the measurement scales on ad liking (Haley & Baldinger, 1991; Walker & Dubitsky, 1994). The measure was a single-item, five-point bipolar measure, with the question to the respondents how much they liked the video, ranging from not at all to a lot.

3.4.4. Control variables

Since the brand that was used in the video was likely not unfamiliar to some of the respondents, it was important to control for brand familiarity. Therefore, a three-item seven-point scale adopted from Kent and Allen (1994) was used in the survey, asking to what extent the respondent was: familiar/unfamiliar, inexperienced/experienced and knowledgeable/not knowledgeable with the brand. Moreover, a question asking if the respondents had ever used the brand themselves was inserted, with a five-point scale (Kent & Allen, 1994) to measure previous experience with the product. Principal component analysis showed that all items were loaded on one factor (Eigenvalue 2.06). The Kaiser-Meyer-Olkin value of .68 ($p < .001$) supported the combination of these items. Together, the items explained 68.7% of the variance in brand familiarity (factor loadings as follows: .78 for Unfamiliar/Familiar, .84 for Inexperienced/Experienced, .86 for Not knowledgeable/Knowledgeable). A reliability test was conducted to confirm internal consistency (Cronbach's α = .77). Finally, the overall score for brand familiarity was created by averaging scores for all three of the sub-scale items. Another control variable was the amount of time the respondents spent watching beauty and/or make-up related content on YouTube. They were asked to answer on a 7-point scale ranging from 'Multiple times a day' to 'Never'. The final control variable in this research tested if the

24

respondent had an interest in beauty or make-up. The respondent was merely able to answer 'Yes' or 'No'.

As mentioned before, YouTube communities often have very loyal subscribers that religiously follow a number of YouTubers. Since the research looks at such a specific topic, namely make-up, it was important to control for the likelihood that the respondent would actually view videos like this when they go on YouTube for themselves. If this was not the case, they would likely not understand the video, or misinterpret it. The effects of the native advertising likely work the best on people who regularly watch these types of videos. Therefore, realism was measured with a statement about the likelihood of the respondents to watch videos like this outside of the context of the experiment: "*This is the kind of video I tend to watch online*". Respondents had to indicate to what extent they agree with the statement. The measurement scale was therefore based on a five-point Likert scale, ranging from 'strongly agree' to 'strongly disagree'.

Table 3.2: Means and standard deviations of all scale variables

Variable	*M*	*SD*
Brand familiarity	-.12	.81
Purchase intention	.004	.92
Source trustworthiness	2.67	.79
Source expertise	2.89	.69
Positive message credibility	3.35	.87
Negative message credibility	4.05	1.34
Video liking	2.82	.91
Brand familiarity	3.79	.82
Frequency of watching YT	3.35	1.95
Interest in beauty/make-up	1.06	.24
Realism	3.81	1.78

3.5. Data Collection

As mentioned before, the research was conducted through an online experimental survey. The survey was developed through the use of Qualtrics. Since the research used a homogenous purposive sampling technique, it aimed to find people that fit the sample description. Therefore, it seemed most logical to start spreading the survey on YouTube itself, by placing a link to the survey with a short description alongside in the comment section of beauty reviews and tutorials of big YouTube make-up gurus like NikkieTutorials, Jeffree Star and MannyMUA. These YouTube channels often have more than a couple million subscribers and often over a million views per video. These videos are basically a place of gathering for those interested in make-up. For the average video, there are about 10,000 to 15,000 comments, and even more people who read them. In the comment section, people are discussing what they think of the video, the YouTuber and the product. Moreover, the survey was posted on Reddit in a couple of subreddits, for example in the subreddit MakeUpAddiction. Reddit is often considered a supportive and close-knit community, where users often reach out to help each other. Lastly, the survey was spread on Facebook, through the use of the personal account of the researcher.

3.6. Manipulation Checks

3.6.1. Stimulus material

In order to ensure that the manipulation of the title and description box in the video was successful, all respondents were asked what product was used in the video. This question was asked after they saw the video, and was asked in order to be able to filter out respondents that did not watch the video 'properly'. Even though the content of the video itself was not of importance to the research, it was to be expected that respondents watched the full video, with or without noticing the title and the description box, and thus also saw what product was used. The answering options were lipstick, mascara, and foundation. These options were chosen to be able to rule out people who did not watch the video even a little bit. These people (n=4) were also removed from the dataset.

3.6.2. Randomization of control variables in conditions

In order to verify that respondents were equally distributed amongst the eight different conditions, manipulation checks had to be conducted for the control variables. The control variables in this research are age, brand familiarity, frequency of watching YouTube, interest in beauty and realism. For all variables, a one-way ANOVA was conducted to compare groups.

A one-way ANOVA revealed that there was not a statistically significant difference in age between the four different description box conditions; $F(3, 254) = .72$, $p = .540$, partial $\eta^2 = .008$. Another one-way ANOVA revealed that there was also not a statistically significant difference in age between the two different title conditions; $F(1, 256) = 2.00$, $p = .158$. Thus, it can be concluded that age was successfully randomized across conditions. The same was done for brand familiarity, where it was revealed that there was not a statistically significant difference in brand familiarity between the four different description box conditions; $F(3, 254) = .85$, $p = .466$, partial $\eta^2 = .010$. Another one-way ANOVA revealed that there was also not a statistically significant difference in brand familiarity between the two different title conditions; $F(1, 256) = .47$, $p = .495$. Thus, it can be concluded that brand familiarity was successfully randomized across conditions.

A third one-way ANOVA revealed that there was not a statistically significant difference in frequency of watching YouTube videos between the four different description box conditions; $F(3, 254) = .83$, $p = .885$, partial $\eta^2 = .003$. This outcome was the same for the frequency of watching YouTube videos in the two title conditions: $F(1, 256) = 2.11$, $p = .147$. Thus, it can again be concluded that frequency of watching YouTube videos was successfully randomized across conditions.

Beauty interest also did not have a statistically significant difference between the four different description box conditions; $F(3, 254) = .85$, $p = .470$, partial $\eta^2 = .010$. This result was the same for the difference between the two different title conditions; $F(1, 256) = .46$, $p = .500$. Thus, it can be concluded that beauty interest was successfully randomized across conditions.

For the final variable, a one-way ANOVA revealed that there was not a statistically significant difference in realism between the four different description box conditions; $F(3, 254) = .60$, $p = .614$, partial $\eta^2 = .007$. Another one-way ANOVA revealed that there was also not a statistically significant difference in realism between the two different title conditions; $F(1, 256) = .18$, $p = .672$. Thus, it can be concluded that realism was successfully randomized across conditions.

3.7. Data Analysis

After a sufficient amount of responses was collected, the Qualtrics survey data was exported into SPSS. SPSS transforms results from Qualtrics into numerical data that is open for

analysis (Babbie, 2007). In order to answer the various hypotheses, the main types of data analyses used in this research were independent samples T-tests and a two-factor ANOVA.

The independent samples T-tests were used to answer hypothesis 1a, 1b, 1c, 1d and 2a. All of these hypotheses tested whether one of the two conditions differed significantly from the other. Independent samples t-tests are commonly used to test the statistical differences between the means of two groups (Pallant, 2013). In the case of this research's hypotheses 1a-1d, every time the mean of one condition's brand attitude change and purchase intention change was compared to that of the other condition.

For hypothesis 2b, which states that "*A video with the title of 'review' will only result in less advertising effectiveness than a video with the title of 'tutorial' if the description box either states that the video is sponsored or contains a coupon code.*", a Multivariate Analysis, or MANOVA was used. The main advantage of MANOVA's are that they are capable of analyzing data and phenomena at an overarching level, since the effects of multiple independent variables on multiple dependent variables can be studied. Moreover, it splits up analysis between each group but also tests possible interaction effects. This can lead to some useful results. As can be read in this hypothesis, the four videos that were titled review are split up in two groups: one group with the sponsored and coupon reviews, and one group with the other two. Since it was necessary to look at the effect of this first group together with the title of the video on advertising effectiveness, it was decided to create a new variable containing these two levels of the description box (sponsored/coupon and empty/not sponsored). This new variable, together with the variable determining the title of the video, were the fixed factors (IVs) in this analysis, against the dependent variables of brand attitude change and purchase intention change.

All main hypotheses were answered with either independent samples t-test or a MANOVA. However, any exploratory analyses that were done to shed more light on why a hypothesis was rejected was conducted through paired-samples t-test, g-power analysis, and one-step multiple regression. Any interesting results that came out of these analyses will be discussed in the discussion section.

3.8. Validity & Reliability

Naturally, it was important to look at the factors affecting the validity and reliability of this research. The validity of the research was influenced by the operationalization of the variables studied. Since the survey included several questions based on scales developed and

tested by other researchers, this would reduce the chances of the measurements not being accurate. However, since the measuring scale for message credibility was only recently developed, this may have caused some discrepancies in terms of the validity of results about message credibility. The reliability of this research was also dependent on measurement scales, but in a different way. For reliability, it was important to make sure that the items that make up one dimension correlated sufficiently. If the same respondent would fill in the same questions, it would have to be so that the same items would give the same answers. If this was not the case, the scales used were not reliable (Santos, 1999). This was tested through the use of a reliability analysis after the data collection in SPSS, looking at the Cronbach's alpha. For this research, an α of .70 and above was considered to be reliable (Salkind, 2011).

4. Results

In this section of the research, the main findings will be presented. Each hypothesis will be discussed separately and will be either accepted or rejected by the end of the paragraph. In order to test the first four hypotheses of this research, which all state something about one of the types of description boxes resulting in less advertising effectiveness than the other, multiple independent t-tests were conducted. Since it was decided that 'advertising effectiveness' will be determined by the two variables of *Brand Attitude Change* and *Purchase Intention Change,* each hypothesis consists of two independent t-tests, one for each variable. Independent t-tests were appropriate analyses for these hypotheses because they compare the differences between two of the four description box conditions for the dependent variables used in this research. The fifth hypothesis (2a) was also tested through the use of an independent samples t-test. The sixth and final hypothesis (2b) was tested by a MANOVA. Below, each paragraph gives the result of each of the separate tests. Moreover, after the rejection or acceptance of the hypothesis, other types of tests will be used to find out what variables could explain the outcome of the main tests. These analyses will be discussed under section 4.4 'Additional and Exploratory Analyses'

4.1. Influence of the Description Box on Advertising Effectiveness

4.1.1. Sponsored vs. not sponsored

Hypothesis 1a assumed that visible sponsorships of a YouTuber in their description box would result in less advertising effectiveness than a visible lack of sponsorships ('This video is NOT sponsored'). An independent samples t-test showed that respondents who saw a 'sponsored' description box ($M = -.17$, $SD = .84$) did not differ significantly in their brand attitude change from respondents who saw a 'non-sponsored' description box ($M = -.21$, $SD = .56$), $t(106.09) = -.35$, *one-tailed p* $= .363$. A second independent samples t-test showed that respondents who saw a 'sponsored' description box ($M = -.16$, $SD = .98$) did not differ significantly in their purchase intention change from respondents who saw a 'non-sponsored' description box ($M = -.13$, $SD = .96$), $t(121) = .17$, *one-tailed p* $= .432$. Since both variables that make up advertising effectiveness do not show a statistically significant difference between each group, it can be concluded that hypothesis 1a is rejected.

4.1.2. Coupon vs. sponsored

Subsequently, an independent samples t-test was conducted to test hypothesis 1b, which assumed that a coupon code used by a YouTuber in their description box would result in less

advertising effectiveness than a visible sponsorship. The test showed that respondents who saw a 'sponsored' description box (M = -.17, SD = .84) did not differ significantly in their brand attitude change from respondents who saw a 'coupon' description box (M = -.04, SD = 1.04), $t(127)$ = .79, *one-tailed p* = .216. A second test revealed that purchase intention change was indeed significantly different for respondents who saw a 'sponsored' description box (M = -.16, SD = .98) than for respondents who saw a 'coupon' description box (M = .15, SD = .74), $t(127)$ = 2.04, *one-tailed p* = .022. Respondents who saw the video where the description box showed a coupon code on average had more purchase intention after they saw the video than before. This stands in contrast to respondents who saw the video where the description box showed that the video was sponsored, who, purely based on the mean, on average had more purchase intention before they saw the video than after. However, a paired-samples t-test would have to prove if this is in fact really true. In summary, purchase intention change for the coupon condition is higher than for the sponsored condition, and thus the disclosure of coupon codes has a higher advertising effectiveness than the disclosure of a sponsorship. This effect is consistent with expectations that were made; sponsored videos turned out to be less effective than coupon videos. Since this difference was only significant for purchase intention change and not for brand attitude change, hypothesis 1b could be partially accepted.

4.1.3. Coupon or sponsored vs. empty

To test hypothesis 1c, another independent samples t-test was conducted. This hypothesis assumed that either a coupon code or a visible sponsorship would result in less advertising effectiveness than an empty description box. The test revealed that respondents who saw a 'sponsored' description box (M = -.17, SD = .84) did not differ significantly in their brand attitude change from respondents who saw an 'empty' description box (M = -.09, SD = .73), $t(158)$ = .53, *one-tailed p* = .298. This was the same for the comparison between the respondents in the 'coupon' description box (M = -.04, SD = 1.04), which proved to not be significantly different in brand attitude change than respondents in the empty condition (M = -.09, SD = .73), $t(133)$ = -.38, *one-tailed p* = .354. The same comparisons were repeated to compare purchase intention change scores for these three conditions. Respondents who saw a 'sponsored' description box (M = -.16, SD = .98) did not differ significantly in their purchase intention change from respondents who saw a 'empty' description box (M = .10, SD = .96), $t(128)$ = .16, *one-tailed p* = .062. This insignificance was similar for the comparison between respondents who

saw a 'coupon' description box (M = .15, SD = .74) respondents who saw an 'empty' description box (M = .10, SD = .96), t(125.79) = -.31, *one-tailed p* = .378. Knowing that for both coupon codes and visible sponsorships, the brand attitude change and purchase intention change did not differ significantly from that of the empty description boxes, hypothesis 1c had to be rejected.

4.1.4. Empty vs. not sponsored

To find out if an empty description box would result in less advertising effectiveness than a visible lack of sponsorships in the description box, which was what hypothesis 1d states, an independent samples t-test showed that respondents who saw a 'non-sponsored' description box (M = -.21, SD = .56) did not differ significantly in their brand attitude change from respondents who saw an 'empty' description box (M = -.09, SD = .73), t(123.79) = 1.04, *one-tailed p* = .150. For purchase intention, this was the same, as respondents who saw a 'non-sponsored' description box (M = -.13, SD = .96) did not differ significantly in their purchase intention change from respondents who saw a 'empty' description box (M = .10, SD = .96), t(127) = 1.38, *one-tailed p* = .085. Having both brand attitude change and purchase intention change not differ amongst the two groups, it had to be concluded that hypothesis 1d was not accepted.

4.2. Influence of the Title on Advertising Effectiveness

Hypothesis 2a assumed that having 'tutorial' as a title would result in less advertising effectiveness than having 'review' as a title. An independent samples t-test showed that respondents who saw a 'tutorial' as a title (M = -.09, SD = .85) did not differ significantly in their brand attitude change from respondents who saw a 'review' as a title (M = -.16, SD = .77), t(256) = -.66, *one-tailed p* = .255. Moreover, respondents who saw a 'tutorial' as a title (M = -.04, SD = .88) proved to not differ significantly in their purchase intention change from respondents who saw a 'review' as a title (M = .05, SD = .96), t(256) = .75, *one-tailed p* = .227. From these results it could be concluded that hypothesis 2a had to be rejected.

4.3. Influence of the Title and Description Box on Advertising Effectiveness

Hypothesis 2b assumed that the only two instances where a review would have less advertising effectiveness than a tutorial would be when the review was sponsored or had a coupon. This also meant that the only two instances where a review would have more advertising effectiveness than a tutorial would be when the review had an empty description box or was not sponsored. As mentioned before, a new variable was created for the description box conditions, where one group contained all responses that were in the coupon or sponsored condition in one

group, and responses that were in the empty or not sponsored condition in the other. This new variable was then combined with the variable that determined the condition based on the title of the video. By performing a MANOVA where these two groups were compared with the two types of titles (review or tutorial), the hypothesis could be accepted or rejected. Accordingly, the MANOVA showed that there was a marginally significant interaction effect between type of title and type of description box for advertising effectiveness, $F(2, 253) = 2.46$, $p = .087$; Wilks' $\Lambda = .98$. It should also be noted that the interaction effect of description box and title of the video on brand attitude change in this MANOVA was marginally significant ($F(1, 254) = 3.74$, $p = .054$; partial $\eta^2 = .015$). Even though theoretically, hypothesis 2b had to be rejected, the marginally significant results do have some meaning behind them. The fact that the p-values are marginal is likely due to the fact that the sample size was relatively small. The marginal significance thus does indicate that the combination of title and description box has some effect on advertising effectiveness.

4.4. Additional and Exploratory Analyses

The previously developed and tested hypotheses were all part of the confirmatory analysis that this research consists of. As mentioned before, it was very important to keep in mind the possible effects of the control variables and other possible influencers on the outcomes of this research. Therefore, this section will discuss the exploratory analyses that were conducted to find out why the main hypotheses were rejected and/or accepted. The variables that will be used for this exploratory analysis are source credibility, message credibility, and video liking, as explained in the theoretical framework.

The sub-variables that will be discussed here are source trustworthiness, source expertise, positive message credibility, negative message credibility and video liking. The control variables that will be discussed here are brand familiarity, realism, frequency of watching YouTube videos, interest in beauty and age.

4.4.1. Changes in attitudes over time

In order to determine if the difference between the means of the pre- and post-test variables for brand attitude and purchase intention is statistically significant, two paired-samples t-tests were run. The first paired-samples t-test was conducted to compare brand attitude before and after the stimulus video across all conditions. The test revealed that there was a statistically significant difference in the scores for pre-vid brand attitude ($M = 5.04$, $SD = 1.09$) and post-vid

brand attitude (M = 4.91, SD = 1.17); $t(257)$ = 2.46, p = .015. The second paired-samples t-tests was conducted to compare purchase intention before and after the stimulus video. The test revealed that there was not a statistically significant difference in the scores for pre-vid purchase intention (M = 4.17, SD = 1.95) and post-vid purchase intention (M = 4.17, SD = 1.90); $t(257)$ = .07, p = .946. These results suggest that across all conditions, people lowered their brand attitude after they saw the video, but in their purchase intention change, they did not have this same effect. In the discussion section of this paper, this result will be further discussed.

4.4.2. Power analyses

Some of the tests that were conducted to find out if the main hypotheses could be rejected or accepted found a nearly statistically significant result. These included hypothesis 1c , where the sponsored condition (M = .16, SD = .98) was compared to the empty condition (M = -.10, SD = .96) for purchase intention change ($t(128)$ = .16, p = .062). Also hypothesis 1d, where the non-sponsored (M = .13, SD = .96) and empty condition (M = -.10, SD = .96) were compared for purchase intention change ($t(127)$ = 1.38, p = .085), revealed an almost-statistically significant effect. In order to find out how large the chance was that there would be a statistically significant effect for these hypotheses, two post hoc power analyses were performed.

The first analysis compared sponsored (n = 62) to empty (n = 68) for purchase intention. The effect size in this study was d = 0.27, which is considered to be a small effect using Cohen's (1988) criteria. The post hoc power analysis revealed that the power to detect a standardized mean difference of d = 0.27 with α = 0.05, using a *one-tailed* t-test when N = 130 is 45.0%. The second analysis compared non-sponsored (n = 61) to empty (n = 68) for purchase intention. The effect size of this study was d = 0.24, which is considered to be a small effect using Cohen's (1988) criteria. The post hoc power analysis revealed that the power to detect a standardized mean difference of d = 0.2 with α = 0.05, using a *one-tailed* test when N = 123 is 37.3%. These results suggests that the comparisons are underpowered, and thus that the chance of there being a statistically significant difference between the sponsored/non-sponsored and empty condition for purchase intention is too small for interpretation. These values thus limit the generalizability of this study's results. In the limitation section, this outcome will be elaborated on.

4.4.3. Explanatory analyses

As mentioned before, it is also important to see if and how variables that were not part of the main hypotheses had an effect on the outcome of the analyses. For both the four description

box conditions as for the two title conditions, regression analyses with the 'remaining' variables were conducted, with brand attitude change and purchase intention change as the dependent variables. All variables were added in the same block. In this section, only the models that turned out to be statistically significant will be discussed. It has to be noted that multicollinearity was tested for in all cases, and all VIF values were below 4, meaning that it can be safely said it was unlikely that effect sizes were depressed because of multicollinearity (O'Brien, 2007).

For the four different description box conditions, a one-step multiple regression was conducted with brand attitude change as the dependent variable. Predictors for all four possible models were source trustworthiness, source expertise, positive message credibility, negative message credibility and video liking. The models for the coupon conditions ($F(5, 61) = 2.68$, $p = .030$, $R^2 = .18$) and the sponsored conditions ($F(5, 56) = 2.69$, $p = .030$, $R^2 = .19$) turned out to be statistically significant. For the coupon conditions, none of the separate predictors within the model were found to be statistically significant. The significance of the entire model is likely due to high interaction effects between the variables, even though alone the variables do not necessarily have an effect on brand attitude change. For the sponsored conditions, *Positive Message Credibility* was found to be a significant, positive and moderate predictor ($\beta = .53$, $p = .014$). Therefore, *Positive Message Credibility* was placed into a regression analysis alone, where it proved that the variable by itself was indeed a statistically significant predictor for brand attitude change in the sponsored conditions ($F(1, 60) = 10.59$, $R^2 = .15$, $\beta = .39$, $p = .002$).

For the four different description box conditions, another one-step multiple regression was conducted with purchase intention change as the dependent variable. Predictors for all four possible models were source trustworthiness, source expertise, positive message credibility, negative message credibility and video liking. The models for the empty conditions ($F(5, 62) = 3.09$, $p = .015$, $R^2 = .20$) and the sponsored conditions ($F(5, 56) = 2.41$, $p = .047$, $R^2 = .18$) turned out to be statistically significant. For the empty conditions, *Positive Message Credibility* was found to be a significant, positive and moderate predictor ($\beta = .47$, $p = .008$). As an additional analysis, *Positive Message Credibility* was placed into a regression analysis alone, where it proved that the variable was indeed a significant predictor for purchase intention change in the empty conditions ($F(1, 66) = 14.07$, $R^2 = .18$, $\beta = .42$, $p < .001$).

For the sponsored condition, *Source Trustworthiness* ($\beta = -.56$, $p = .014$), *Source Expertise* ($\beta = .34$, $p = .015$) and *Positive Message Credibility* ($\beta = .51$, $p = .018$). were found to

be the statistically significant predictors. To see if these variables were also of influence by themselves, all three predictors were also placed into regression analyses alone. *Source Trustworthiness* alone was not a significant predictor for purchase intention change in the sponsored conditions ($F(1, 60) = .08$, $R^2 = .001$, $\beta = -.04$ $p = .781$). *Source Expertise* was the only variable that, by itself, was significant predictor for purchase intention change in the sponsored conditions ($F(1, 60) = 4.43$, $R^2 = .07$, $\beta = .26$ $p = .040$). *Positive Message Credibility* alone was also not a significant predictor for purchase intention change in the sponsored conditions ($F(1, 60) = 1.14$, $R^2 = .019$ $\beta = .14$, $p = .289$).

For the two different title conditions, a one-step multiple regression was conducted with brand attitude change as the dependent variable. Predictors for the two possible models were source trustworthiness, source expertise, positive message credibility, negative message credibility and video liking. The models for both the review conditions ($F(5, 118) = 3.96$, $p = .002$, $R^2 = .14$) and the tutorial conditions ($F(5, 128) = 5.61$, $p < .001$, $R^2 = .18$) turned out to be statistically significant. For the review condition, *Positive Message Credibility* was found to be the only statistically significant predictor ($\beta = .38$, $p = .004$). As an additional analysis, *Positive Message Credibility* was placed into a regression analysis alone, where it proved that the variable by itself was indeed a statistically significant predictor for brand attitude change in the review conditions ($F(1, 122) = 13.8$, $R^2 = .10$, $\beta = .32$, $p < .001$).

For the tutorial conditions, *Source Trustworthiness* ($\beta = .39$, $p = .001$), *Source Expertise* ($\beta = .32$, $p = .002$) were found to be statistically significant predictors. As an additional analysis, both predictors were also placed into a regression analysis alone. *Source Trustworthiness* alone proved to be a statistically significant predictor for brand attitude change in the tutorial conditions ($F(1, 132) = 16.1$, $R^2 = .11$, $\beta = .19$ $p < .001$). *Source Expertise* alone also was statistically significant predictor for brand attitude change in the tutorial conditions ($F(1, 132) = 13.2$, $R^2 = .09$, $\beta = .30$, $p < .001$).

For the two different title conditions, another one-step multiple regression was conducted with purchase intention change as the dependent variable. Predictors for the two possible models were source trustworthiness, source expertise, positive message credibility, negative message credibility and video liking. Neither the review nor the tutorial conditions turned out to have a statistically significant model with the combination of the variables mentioned before. However, when looking at the coefficients table, we can see that *Source Expertise* is a statistically

significant variable for the tutorial conditions model. Therefore, *Source Expertise* alone was placed into a regression analysis. The variable indeed was found to be a statistically significant predictor for purchase intention change in the tutorial conditions ($F(1, 132) = 10.92$, $R^2 = .08$, β = .28, $p = .001$).

4.4.4. Analyses of control variables

As there were many control variables present in this study, it was very important to keep in mind the possible effects that these variables may have had on advertising effectiveness. The five control variables in this study were age, frequency of watching YouTube, interest in beauty, brand familiarity and realism. By using a simple linear regression analysis, the control variables were be tested for their possible effects on brand attitude change and purchase intention change.

A one-step multiple regression was conducted with brand attitude change as the dependent variable. Predictors for the possible model were age, frequency of watching YouTube, interest in beauty, brand familiarity and realism. The model ($F(5, 224) = 1.42$, $p = .217$, $R^2 = .031$) turned out to not be statistically significant. However, *Realism* was found to be a statistically significant predictor (β = -.21, $p = .021$) within the model, as the only variable. As an additional analysis, *Realism* was placed into a regression analysis alone, where it proved that the variable by itself was indeed a statistically significant predictor for overall brand attitude change ($F(1, 256) = 4.72$, $R^2 = .018$, β = -.14, $p = .031$).

A second one-step multiple regression was conducted with purchase intention change as the dependent variable. Predictors were age, frequency of watching YouTube, interest in beauty, brand familiarity and realism. The model ($F(5, 224) = 1.20$, $p = .310$, $R^2 = .026$) turned out to not be statistically significant. However, *Realism* was again found to be a weak negative significant predictor (β = -.18, $p = .049$), but only when taken together with the other variables. By itself, *Realism* was not a significant predictor for purchase intention change. ($F(1, 256) = 1.00$, $R^2 = .004$, β = -.062, $p = .319$).

4.4.5. Additional ANOVAs

In order to see whether there were any differences between conditions with regards to the subvariables (source trustworthiness, source expertise, positive message credibility, negative message credibility, video liking). Therefore, two ANOVAs were conducted to see if these variables had a different effect in the various conditions.

A one-way ANOVA revealed that there was not a statistically significant difference in source trustworthiness ($F(3, 254) = 1.19$, $p = .313$, partial $\eta^2 = .014$), source expertise ($F(3, 254) = .99$, $p = .397$, partial $\eta^2 = .012$), positive message credibility ($F(3, 254) = .20$, $p = .897$, partial $\eta^2 = .002$) negative message credibility ($F(3, 254) = 2.55$, $p = .056$, partial $\eta^2 = .029$) and video liking ($F(3, 254) = .29$, $p = .834$, partial $\eta^2 = .003$) between the four different description box conditions.

A second one-way ANOVA revealed that there was not a statistically significant difference in source trustworthiness ($F(1, 256) = .33$, $p = .567$, partial $\eta^2 = .001$), source expertise ($F(1, 256) = .41$, $p = .523$, partial $\eta^2 = .002$), positive message credibility ($F(1, 256) = .11$, $p = .739$, partial $\eta^2 < .001$) negative message credibility ($F(1, 256) = .19$, $p = .661$, partial $\eta^2 = .001$) and video liking ($F(1, 256) = 1.02$, $p = .313$, partial $\eta^2 = .004$) between the two different title conditions. Overall, respondents seemed to be largely unaffected by the title and description box of each video on attitude outside of hypotheses.

.

5. Discussion & Conclusion

In the final section of this research, the findings of the analyses conducted are discussed. This study focused on the effect of different types of description box contents and titles in relation to sponsorships and coupon codes on the advertising effectiveness of a YouTube videos. More specifically, the research focused on the beauty industry and the way it functions on YouTube. After conducting an online experiment, the brand attitude change and purchase intention change were examined using a quantitative approach. The aim of this research was phrased through the following research question: *How does the context of YouTube videos affect the effectiveness of advertising?*. With six hypotheses that were either accepted or rejected, this section of the study will connect the results of the analyses with literature that was used in the theoretical framework. Moreover, the many implications that this research has, which are both theoretical, societal and business-related, will be elaborated on. Lastly, the limitations of the study and suggestions for future research will be discussed.

5.1. Discussion

The first four hypotheses of this study concerned the effects of different types of description boxes on the advertising effectiveness of the video used in the online experiment. The four types of description boxes were: (1) coupon, (2) empty, (3) not sponsored, and (4) sponsored. As discussed in previous literature, it was to be expected that people would be less affected by the advertising in the video if it was disclosed that somehow the YouTuber would be paid for it or had gotten the product for free (Wu, 2016; Wojdynski & Evans, 2016; Carlson, 2015). However, analyses showed that this was not the case for three out of the four comparisons that were done, and for the fourth comparison the effect was opposite of what was expected. It can therefore be concluded that different description boxes do not change how people perceive or look at the video. These description boxes were expected to influence respondents in a way that whenever a sponsorship was disclosed, the viewer would alter its perception of the brand in the video by having a more negative attitude towards it and being less willing to buy its products (Boerman et al., 2014). However, it turned out that these effects held off. This was also again proven through the G-power analyses that were conducted in the additional analyses section. Rather, the participants in the experiment did not seem to 'notice' the disclosures or the absence thereof, which is in line with the expectations of Hoofnagle and Meleshinsky (2015), who wrote that often enough, people do not recognize the disclosure of sponsorships. This fact was also

confirmed through the study of Wojdynski and Evans (2015), who in their research found that only 8% of their participants (N=242) recognized native advertising as true 'advertising content'. The authors also stated that when advertising is integrated into the content, as is often the case in these types of videos, consumers experience the advertisements as less annoying and disruptive. Cho & Cheon (2004) add to this by stating that native advertising can foster positive reactions from consumers for the exact same reason. What is important to mention too, is that literature concerning the disclosure of native advertising and its effectiveness is contradictory. This is another possible reason for the contradicting results in this research. For example, scholars and governmental institutions like the Federal Trade Commission have not agreed on what is the best possible placement for a native advertisement for consumers to be able to recognize it better (Wojdynski & Evans, 2015; Federal Trade Commision, 2015).

It has to be noted that the description box of YouTube videos is usually hidden from the viewer until they click the button to reveal the description box. However, in this study, the description box was deliberately shown to the viewer. This suggests that since there is no effect when the description box is shown, the effect will be even smaller when respondents would have to click a button to reveal the description box. In that case, it merely is the curiosity of the respondents for the contents of the description box that determines whether or not the sponsorship or lack of sponsorship is disclosed or not. In this way, this research showed that viewers overall are not perceptive of the contents that are not part of the actual video, in this case the description box, even if this content is put right in front of them to see. This implies that viewers can easily be persuaded by an advertisement or sponsorship since they are not aware that it is happening. This research thus showed that native advertising like this can be extremely deceiving. This is where the problem of ethics comes in. Schauster, Ferrucci and Neill (2016) conducted in-depth interviews with digital journalists and marketing-communication professionals to find out what their perceptions of native advertising were, and how it affected their conscience of social responsibility. Not only did they find it deceptive, but also they found it unethical because consumers cannot find out easily what is sponsored content and what is not.

The additional paired samples t-tests that were conducted showed that across all conditions, people lowered their brand attitude after they saw the video. It has to be noted that Britt, the YouTuber, was very positive about the Maybelline foundation in her video. These two facts together make for an interesting insight; could the positivity in the video have been

interpreted as persuasion by an ad, no matter what the description box or title said? Possibly, the respondents did perform some form of self-regulation, by showing that they were resistant to the 'persuasion' present in the video and reporting a lower brand attitude after the video than before (Wheeler, Briñol & Hermann, 2007). If that is indeed the case, then the video was not deceptive at all, and respondents assumed that even though the description box was empty or said that the video was not sponsored, there was some type of payment or deal between the YouTuber and Maybelline. The content of the video was then more important than what was in the description box. It could also be a combination of this and the fact that some respondents may not have seen what was in the description box at all. Moreover, the analyses on the sub-variables showed that these variables did not differ amongst the various conditions. This means that, for example, source credibility was rated similarly amongst respondents who got to see the sponsored video and the non-sponsored video. Obviously, this says something about the effect of the disclosure of sponsorships. It means that the conditions that respondents were put in had no influence on their interpretation of the video. Since the source credibility scores were both relatively low, it can be assumed that respondents found the positive 'review' of the product in the video not so believable in any case. This was confirmed by some comments of respondents who had to indicate why they were was not able to watch the whole video, and were removed from the valid sample:

> *"It seems too exaggerated to me, like I don't really like it when people talk about brands like that. It seems a little too sponsored even though it says that it's not sponsored."*
> *"Too positive, no critical questions. If a product is good, she should of course let the viewer know this but this made it slightly unbelievable."*

All of these results suggest that indeed, the consumers either did not care about or did not see the contents of the description box, and based their answers in the survey solely on what they saw in the video.

Hypothesis 2a and 2b both stated some expectations with regards to the effect of a certain title on the effectiveness of advertising within the video. Both hypotheses were rejected, and so it has to be contemplated as to why these expectations were not correct after all. As became clear from the theoretical framework, there is not much literature to be found on these effects This is mostly because the notions of 'review' and 'tutorial' are relatively new and are also very specific to the YouTube community. The expectations that the hypotheses were based on came from the

assumption that is made when you hear the word 'review': a review about a concert, a restaurant, or a tech product that are presented in a newspaper or on a website are considered to be honest, truthful representations of what a certain person (a journalist, someone who ate at a restaurant or someone who bought a new TV) thinks of the product or service. The assumptions of honesty and/or truthfulness hold because the person writing the review does not have any interest in writing the review other than 'helping' other consumers to make the right decision. This is where a review of a restaurant, for example, and a review of a make-up product by a big YouTuber are different: YouTubers have a commercial interest just like reviewers that work for newspapers and magazines, but the Youtuber cannot 'lose' their job is they do not produce good work. Moreover, journalists are expected to exert a level of objectivity, as this is what ethical reporting is based on. This is not necessarily required from YouTubers. Possibly, viewers could already assume that, even though a video says it is a review, the YouTuber is out for money and thus the title becomes unimportant. This could be a reason why tutorials and reviews are not considered to be different, even under different description box conditions that could contradict the notion of a 'review' (a sponsored review is likely highly unbelievable). It may be concluded that rather than the title being a central cue (Petty, Cacioppo & Schuurman, 1983) in the process towards persuasion by an advertisement, it rather is a peripheral cue; the title is only understood in context of the rest of the information that is being given.

The exploratory analyses of the sub-variables' effects on advertising effectiveness present some interesting results. In the sponsored conditions, message credibility and source credibility influence the respondents' scores for one of the variables on advertising effectiveness. This suggests that the disclosure of the sponsorship did affect how a respondent saw the YouTuber with regards to credibility and subsequently this affects advertising effectiveness. However, since the same is the case in the empty conditions, where message credibility has a significant effect on purchase intention change, these assumptions cannot form a concrete pattern of results that supports the theory presented before. The analyses with the sub-variables on the title conditions suggest that in the tutorial conditions, source credibility is an important predictor for advertising effectiveness, and it is not within the review condition. This is another interesting finding that contradicts expectations, as the notion of 'review' would often be considered linked to how credible a source is: a review is only valuable if the source it comes from is credible.

5.2. Implications of the research

The question remains; should YouTubers be even more explicit in their disclosure of advertisements and sponsorships? Or are they already doing their utmost by disclosing it in the description box and is the recognition of this in the hands of the consumer? The developers of the Dutch 'Social Code' would say so (https://www.desocialcode.nl/). They established these guidelines to disclose sponsorships in the description box because they wanted to be honest and open with their viewers, but they are also still a group of YouTubers that wants to earn money and not have their loyal followers stop viewing and get angry because they get paid to talk positively about all of these things. However, it is still true that the disclosure of advertisements and sponsorships could be a lot more explicit than putting it in the description box: mentioning it in the video itself for example, or putting a still in the video for a view seconds disclosing the sponsorship. In that way, viewers have no way around the disclosure and are basically forced to live with this reality and make their own choice. Maybe they don't care and just continue watching, maybe they do and they stop the video. This would be the most ethically approved way to handle these situations. However, a lot of YouTubers often make a living off of their videos and thus they have to gain income and be able to grow. The credibility of the YouTuber will likely be compromised because they are forced to disclose all of the times they are being paid to talk about something (Dekker & van Reijmersdal, 2013). Of course, hiding sponsorships and advertisements continuously is not an ethical way to work.

The problem of ethics, in YouTube videos but more generally in all platforms and content that use native advertising, is something that should be brought forward more prominently by governments, but possibly also by YouTube itself. Even though YouTube is merely the medium through with sponsorships and native advertisements are fostered, they should be the instigator for YouTube becoming a more consumer-friendly platform, thus meaning being more honest and open with consumers about the nature of certain videos. For example, YouTube could create a simple labelling system (Hyman, Franklyn, Yee & Rahmati, 2017), where YouTubers have to 'tick a box' if they are being sponsored before they upload the video. Moreover, it is the job of the government to regulate ethics within all scopes of the media, and therefore it should also act for consumers on YouTube. Even though many of these regulations already exist, they are not executed well. For example, the European Union has established a report and guidelines for parties with a commercial interest on how to deal with

native advertising. One of the laws under the Unfair Commercial Practices Directive states that 'using editorial content in the media to promote a product where a trader has paid for the promotion without making that clear in the content or by images or sounds clearly identifiable by the consumer' is considered unfair and can be punished on the grounds of this (European Parliament, 2005). However, this is exactly the problem with native advertising on YouTube and still the reality remains that a vast amount of YouTubers does not adhere to these guidelines. Likewise, the USA Federal Trade Commission has set up guidelines and laws that are similar if not identical to those of the EU, and the same happens. By establishing more thorough and clear rules and regulations for native advertising and sponsorships, the government can set certain standards for YouTube as an industry specifically, and other advertisers alike, to adhere to, with risk of a fine if they fail to adhere to it. There are so many factors that influence the level of persuasion knowledge a consumer can have (Ham, Nelson & Das, 2015), and therefore Wojdynski and Evans (2018) have made a case for a scale to measure sponsorship transparency, which could be used by regulatory institutions to, for example, measure and determine if a YouTuber is being transparent enough.

5.3. Limitations

As any other study, this research has some limitations that need to be addressed. The most pressing limitation of the study is the limited sample size. The power analysis that was conducted revealed that the comparisons between different conditions were underpowered, meaning that the chance that there would be a significant difference between groups in advertising effectiveness was very small. However, it has to be noted that the effect size, which was very small in this case and determined whether or not the comparison was underpowered, is heavily influenced by the sample size. If the sample size would have been bigger, the effect size and thus the power of the comparisons would have been higher, meaning that hypothesis 1c and 1d could have possibly been accepted. It also has to be noted that the small effect size that was retrieved here possibly says something about the real-world significance of the comparisons between different description boxes; the disclosure of a sponsorship just does not really influence advertising effectiveness, as was stated before.

Another limitation is the fact that the study was based on an online experiment. The researcher did not have control over the experimental situation as it was conducted online, and not in a laboratory setting (Reips, 2000). It is possible that the participants did not watch the

video with full attention or watch the video at all, or that a participant filled out the survey more than once. These are common issues for online experiments, but it also important to note that online experiments have some advantages over traditional laboratory experiments. They usually have higher external validity because participants can participate at their own time, in a for-them comfortable setting (Crano, Brewer & Lac, 2015; Reips, 2000).

Another limitation of this study concerns the sampling method. The sampling method was homogenous purposive sampling, mainly because respondents had to have two main characteristics: they had to be female and between the ages of 16 and 55. However, it was also a combination of convenience sampling and snowball sampling, as the link to the survey was posted on social media and it was asked if respondents wanted to share the link with their friends. Convenience sampling has many issues of representativeness, as there is a higher likelihood that the sample gets shared with people that are all fairly similar (same country, same socio-economic status, etc.), similar to the representativeness issues with snowball sampling due to the unequal chances of people within the target group to be reached for the survey (Fricker, 2008; Grossnickle & Raskin, 2001). However, the choice for these sampling methods are justified because they have in the end helped to obtain the right participants through various social media (Oosterveer, 2014).

A third limitation revolves around the stimulus material and the way in which it was presented in the online survey. First of all, the YouTuber in the video, BasicallyBritt, was likely unknown to each of the participants, as she 'only' has 7,000 followers, which is relatively small in the YouTube community. The big Youtube make-up gurus often have a couple million followers and thus it can be considered more logical that they would get paid for a video as opposed to Britt. A commenter on Reddit even made a comment saying that the researcher should 'try using a YouTuber with a larger subscription base (1-5M subscribers)'. Another person who had to indicate why she did not watch the whole video stated:

"If I had seen more of her videos then I might've known that she is pretty critical about stuff."

Using an existing video (thus, not intended for research purposes) by a big YouTuber would have added more realism to the survey, but it was chosen not to as it was considered important to control for the content of the video. The additional regression analyses that were conducted on the control variables proved this, as realism was found to have a significantly negative effect on brand attitude change. Moreover, as the video was merely uploaded for this research, it did not

have any views nor likes. This was again also noted by the commentor on Reddit, showing that it did not go unnoticed and was likely also noticed by other respondents. This is another aspect that could have slightly decreased the believability of the video.

5.4. Suggestions for future research

The outcomes of this study foster many possibilities for future research to build on and around the topic of YouTube, native advertising, and disclosure. First of all, it would be interesting to see if disclosure of a sponsorship or lack thereof within the video itself does have a significant effect of advertising effectiveness, as compared to disclosure in the description box. In this research, the disclosure was quite subtle and maybe even hard to notice; possibly a disclosure within the video itself, which the respondents cannot get around, would generate a bigger effect. Again, it is hard to say if the advertising effectiveness goes up or down because of this, as scholars are not yet sure whether viewers appreciate the honesty of the disclosure or feel like they are being betrayed because of the sponsorship.

Future research could also go beyond the make-up industries and look at other types of reviews and/or videos in which a YouTuber (possibly) gets paid to talk about a product. For example, there are tons of YouTube channels that dedicate their time to unboxing all sorts of tech products, as well as channels that test out new video games. This also makes for another interesting possibility for research, as it could be interesting to see if men react differently to the disclosure of sponsorships than women. Of course, the research concerning native advertising could also extend itself towards Instagram posts, even though here it is often more clear if someone is being sponsored because of the 'tags'. However, it would be interesting to see how these sponsored posts are perceived in relation to message and source credibility.

The findings of the additional exploratory analyses indicated that in various conditions, there were different sub-variables that had different types of effects on advertising effectiveness. For example, in the sponsored condition, positive message credibility negatively influenced brand attitude change, whilst in the empty condition it positively influenced purchase intention change. The explanation of these effects was not investigated in this research, but could provide for some interesting studies in the future. More specifically, future research could analyze if and how viewers of YouTube videos link a creator to the brand that they are talking about and/or are promoting. As discussed before, the study found some interesting but confusing results on the effects of the sub-variables on advertising effectiveness. Since the sub-variables that were chosen

were all expected to have an effect on advertising effectiveness according to literature, future research could investigate the strength of these relationships again. Moreover, this research looked at brand attitude and purchase intention as measures for advertising effectiveness. Future research could attempt to look at other measures for advertising effectiveness, for example through brand awareness, retention and consumer engagement.

5.5 Conclusion

This study investigated the ways in which disclosure of sponsorships on YouTube influenced advertising effectiveness. The results showed, in contrast to some other similar studies, that disclosure of sponsorships does not have an effect on brand attitude and purchase intention. Since no other conclusive evidence was found as to why this effect stayed out, it can be carefully stated that this may have to do with the fact that consumers do not notice the subtle disclosures that are often 'hidden' in the description boxes of YouTube videos. Even though governments and YouTubers alike, like those of De Social Code, have agreed that there are some ethical issues that come with the deceptive nature of (undisclosed) sponsorships, this study showed that disclosures in description boxes alone are not enough. The difficulty that comes with more explicit disclosures, for example by verbally stating that the YouTuber is indeed sponsored to talk about a certain product, is that according to some studies it may negatively affect the brand attitude and credibility of the YouTuber him or herself. However, scholars are still not definitive as to the true extent of these effects, and often different studies state opposite things. This shows that the topic of sponsorship disclosure needs to be studied more carefully. Having a more clear view of consumers' behaviors and opinions in relation to disclosures can help governments, companies and YouTubers to establish a more fair and ethical playing field that can still serve commercial interests.

References

Appelman, A., & Sundar, S. S. (2016). Measuring message credibility: Construction and
validation of an exclusive scale. *Journalism & Mass Communication Quarterly*, *93*(1),
59-79. doi: 10.1177/1077699015606057

Aqsa, M., & Kartini, D. (2015). Impact of online advertising on consumer attitudes and interests
buy online (survey on students of internet users in Makassar). *International Journal of
Scientific & Technology Research*, *4*(4), 230-236.

Babbie, E. (2007). *Introduction to social research*. (4th edition). Stamford, CT: Cengage
Learning Inc.

Bandura, A. (1986). *Social foundations of thought and action: A social cognitive theory*.
Englewood Cliffs, NJ, US: Prentice-Hall, Inc.

Biel, A. L., & Bridgwater, C. A. (1990). Attributes of likable television commericals. *Journal of
Advertising Research, 30*(3), 38-44.

Boerman, S. C., Reijmersdal, E. A., & Neijens, P. C. (2014). Effects of sponsorship disclosure
timing on the processing of sponsored content: A study on the effectiveness of European
disclosure regulations. *Psychology & Marketing*, *31*(3), 214-224.
doi: 10.1002/mar.20688

Bryant, S. (2004). *Computer-based experiments: Obstacles*. United States: University of
South Florida.

Carlson, M. (2015). When news sites go native: Redefining the advertising–editorial divide in
response to native advertising. *Journalism, 16*(7), 849-865.
doi: 10.1177/1464884914545441

Chang, S., Molesworth, M., & Grigore, G. (2015). YouTube Beauty Brands and the Dynamic
Processes of Prosumption. *Proceedings Of The 40th Annual Macromarketing
Conference*. *Chicago, Illinois. June 25-28th 2015*.

Chen, H., Chiang, R. H., & Storey, V. C. (2012). Business intelligence and analytics: from big
data to big impact. *MIS Quarterly*, *36(*4), 1165-1188.

Cho, C., & Cheon, H. (2004). Why do people avoid advertising on the Internet? *Journal of
Advertising, 33*(4), 89-97. doi: 10.1080/00913367.2004.10639175

Cialdini, R. B., Petty, R. E., & Cacioppo, J. T. (1981). Attitude and attitude change. *Annual
Review of Psychology*, *32*(1), 357-404. doi: 10.1146/annurev.ps.32.020181.002041

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences.* Hillsdale, NJ: Prentice-Hall.

Commissariaat voor de Media. (2017). *Hoe kan de transparantie van reclame in online video's vergroot worden voor minderjarigen?*. Retrieved from https://www.cvdm.nl/wp-content/uploads/2017/11/Literatuuronderzoek-Transparantie-Rozendaal-Van-Reijmersdal.pdf

Crain, M. (2014). Financial markets and online advertising: reevaluating the dotcom investment bubble. *Information, Communication & Society*, *17*(3), 371-384.
doi: 10.1080/1369118X.2013.869615

Crano, W. D., Brewer, M. B., & Lac, A. (2015). *Principles and methods of social research*. New York, NY: Routledge.

Dekker, K., & Van Reijmersdal, E. A. (2013). Disclosing celebrity endorsement in a television program to mitigate persuasion: How disclosure type and celebrity credibility interact. *Journal of Promotion Management*, *19*(2), 224-240.
doi: 10.1080/10496491.2013.769473

Digital Marketing Education. (2018). *Youtube Statistics - 2018*. Retrieved from https://fortunelords.com/youtube-statistics/

DiNucci, D. (1999). Fragmented future. *Print*, *53*(4), 32-33.

Ebizmba.com. (2017). *Top 15 Most Popular Video Websites*. Retrieved from http://www.ebizmba.com/articles/video-websites.

Eliashberg, J., & Shugan, S. M. (1997). Film critics: Influencers or predictors?. *Journal of Marketing, 61*(2), 68-78. doi: 10.2307/1251831

Ellison, N. B. (2007). Social network sites: Definition, history, and scholarship. *Journal of Computer-Mediated Communication*, *13*(1), 210-230.
doi: 10.1111/j.1083-6101.2007.00393.x

ESHCC. (2017). *Methodological Guidelines Thesis Research*. Rotterdam, the Netherlands: Erasmus University Rotterdam.

Etikan, I., Musa, S. A., & Alkassim, R. S. (2016). Comparison of convenience sampling and purposive sampling. *American Journal of Theoretical and Applied Statistics*, *5*(1), 1-4.
doi: 10.11648/j.ajtas.20160501.11

Euromonitor. (2016). *Shiseido Annual Report 2015: Basic Data Section* (p. 76). Shiseido.

European Parliament (2005). *Unfair Commercial Practices Directive*. Strasbourg: Office for
Official Publications of the European Union. Retrieved from https://eur-
lex.europa.eu/legal-content/EN/TXT/?uri=celex%3A32005L0029

Federal Trade Commission. (2015). Native advertising: A guide for businesses.
Retrieved from https://www.ftc.gov/tips-advice/business-
center/guidance/nativeadvertising-guide-businesses

Fishbein, M., & Ajzen, I. (1977). *Belief, attitude, intention, and behavior: An introduction to
theory and research*. Reading, MA: Addison-Wesley

Fricker, R. D. (2008). Sampling methods for web and e-mail surveys. In Fielding, N., Lee, R. M.,
& Blank, G. (Eds.). *The SAGE handbook of online research methods*: SAGE
Publications Ltd

García-Rapp, F., & Roca-Cuberes, C. (2017). Being an online celebrity: Norms and expectations
of YouTube's beauty community. *First Monday, 22*(7). doi: 10.5210/fm.v22i7.7788

Gerhards, C. (2017). Product placement on YouTube: An explorative study on YouTube
creators' experiences with advertisers. *Convergence*. Advance online publication
doi: 10.1177/1354856517736977

Gill, P., Arlitt, M., Li, Z., & Mahanti, A. (2007). Youtube traffic characterization: a view from
the edge. In *Proceedings of the 7th ACM SIGCOMM conference on Internet
measurement. Palo Alto, California. June 12th, 2007* . doi: 10.1145/1298306.1298310

Goldberg, M. E. (1990). A quasi-experiment assessing the effectiveness of TV advertising
directed to children. *Journal of Marketing Research*, *27*(4) 445-454. doi:
10.2307/3172629

Goulston, M., & Ullmen, J. (2013). *Real influence: Persuade without pushing and gain without
giving in*. AMACOM Div American Mgmt Assn.

Grossnickle, J., and Raskin, O. (2001). *The Handbook of Online Marketing Research*. New
York, NY: The McGraw-Hill Companies, Inc.

Gurău, C. (2008). Integrated online marketing communication: implementation and
management. *Journal of Communication Management*, *12*(2), 169-184.
doi: 10.1108/13632540810881974

Haley, R. I., & Baldinger, A. L. (1991). The ARF copy research validity project. *Journal of
Advertising Research*. *31*(2), 11-32.

Ham, C. D., Nelson, M. R., & Das, S. (2015). How to measure persuasion knowledge. *International Journal of Advertising*, *34*(1), 17-53. doi: 10.1080/02650487.2014.994730

Hoofnagle, C., & Meleshinsky, E. (2015). Native advertising and endorsement: Schema, source-based misleadingness, and omission of material facts. *Technology Science*.

Hovland, C. I., & Weiss, W. (1951). The influence of source credibility on communication effectiveness. *Public Opinion Quarterly*, *15*(4), 635-650. doi: 10.1086/266350

Hudson, S., & Hudson, D. (2006). Branded entertainment: a new advertising technique or product placement in disguise?. *Journal of Marketing Management, 22*(5-6), 489-504. doi: 10.1362/026725706777978703

Hyman, D. A., Franklyn, D., Yee, C., & Rahmati, M. (2017). Going native: Can consumers recognize native advertising: Does it matter. *Yale Journal of Law & Technology.*, *19*(77), 79-112.

Ikonen, P., Luoma-aho, V., & Bowen, S. A. (2017). Transparency for sponsored content: analysing codes of ethics in Public Relations, Marketing, Advertising and Journalism. *International Journal of Strategic Communication, 11*(2), 165-178.

Jaakonmäki, R., Müller, O., & vom Brocke, J. (2017, January). The impact of content, context, and creator on user engagement in social media marketing. In *Proceedings of the 50th Hawaii International Conference on System Sciences. Waikoloa, Hawaii. January 4-7, 2017*.

Kaplan, A. M., & Haenlein, M. (2010). Users of the world, unite! The challenges and opportunities of Social Media. *Business Horizons*, *53*(1), 59-68. doi: 10.1016/j.bushor.2009.09.003

Kent, R. J., & Allen, C. T. (1994). Competitive interference effects in consumer memory for advertising: the role of brand familiarity. *Journal of Marketing*, *58*(3), 97-105.

Kim, H. S., & Lynn Damhorst, M. (1999). Environmental attitude and commitment in relation to ad message credibility. *Journal of Fashion Marketing and Management: An International Journal, 3*(1), 18-30. doi: 10.1108/eb022545

Ko, E., Kim, E., Taylor, C. R., Hoon Kim, K., & Jeong Kang, I. (2007). Cross-national market segmentation in the fashion industry: A study of European, Korean, and US consumers. *International Marketing Review, 24*(5), 629-651. doi: 10.1108/02651330710828022

Lafferty, B. A., Goldsmith, R. E., & Newell, S. J. (2002). The dual credibility model: The
influence of corporate and endorser credibility on attitudes and purchase intentions.
Journal of Marketing Theory and Practice, 10(3), 1-11.
doi: 10.1080/10696679.2002.11501916

Lemanski, J. L., & Hyung-Seok, L. (2012). Attitude certainty and resistance to persuasion:
Investigating the impact of source trustworthiness in advertising. *International Journal
of Business and Social Science*, *3*(1), 66-75.

Likert, R. (1932). A Technique for the Measurement of Attitudes. *Archives of Psychology*,
22(140), 1–55.

L'Oréal. (2017). *Annual Report 2016* (p. 14). L'Oréal.

Lutz, R. J., MacKenzie, S. B., & Belch, G. E. (1983). Attitude toward the ad as a mediator of
advertising effectiveness: Determinants and consequences. *ACR North American
Advances*.

McGraw, K. O., Tew, M. D., & Williams, J. E. (2000). The integrity of web-delivered
experiments: Can you trust the data*? Psychological Science, 11*(6), 502–506.
doi: 10.1111/1467-9280.00296

Meindersma, C. (2017). Nieuwe regels voor YouTubers, de Social Code. Retrieved from
https://www.marketingfacts.nl/berichten/nieuwe-regels-voor-youtubers-de-social-code

Miller, N. E., & Dollard, J. (1941). *Social Learning and Imitation*. New Haven, US: Yale U.
Press.

Muijs, D. (2010). *Doing quantitative research in education with SPSS.* Thousand Oaks, CA:
Sage.

Nebenzhal, I. D., & Jaffe, E. D. (1998). Ethical dimensioins of advertising executions. *Journal of
Business Ethics, 17*(7), 805-815.

O'Brien, R.M. (2007). A caution regarding the rules of thumb for variance inflation factors.
Quality & Quantity, *41*(5), 673-690. doi: 10.1007/s11135-006-9018-6.

O'Reilly, T. (2007). What is web 2.0: Design patterns and business models for the next
generation of software. *International Journal of Digital Economics* (65), 17-37.

Ohanian, R. (1990). Construction and validation of a scale to measure celebrity endorsers' perceived expertise, trustworthiness, and attractiveness. *Journal of Advertising*, *19*(3), 39-52. doi: 10.1080/00913367.1990.10673191

Oosterveer, D. (2014, January 27). Social media in Nederland 2014: Jongeren blijven Facebook trouw. Retrieved from http://www.marketingfacts.nl/berichten/nationale-social-mediaonderzoek-2014

Pallant, J. (2007a). Factor analysis. In Pallant (Ed.), SPSS survival manual. *A step by step guide to data analysis using IBM SPSS* (pp. 179-199). New York, NY: Mc-Graw Hill.

Pallant, J. (2007b). Checking the reliability of a scale. In Pallant (Ed.), SPSS survival manual. *A step by step guide to data analysis using IBM SPSS* (pp. 101-105). New York, NY: Mc-Graw Hill.

Perloff, R. M. (2014). Social media effects on young women's body image concerns: Theoretical perspectives and an agenda for research. *Sex Roles*, *71*(11-12), 363-377. doi: 10.1007/s11199-014-0384-6

Petty, R. E., Cacioppo, J. T., & Schumann, D. (1983). Central and peripheral routes to advertising effectiveness: The moderating role of involvement. *Journal of Consumer Research*, *10*(2), 135- 146. doi: 10.1086/208954

Petty, R. E., & Cacioppo, J. T. (1986). The elaboration likelihood model of persuasion. *Advances in Experimental Social Psychology*, *19*, 123-205. doi: 10.1016/S0065-2601(08)60214-2

Pixability. (2015). *Beauty on YouTube* (pp. 10-32). Pixability.

Pixability. (2016). Women's Wear Daily.

Reino, S., Hay, B. (2011). The Use of YouTube as a Tourism Marketing Tool, 42nd Annual Travel & Tourism Research Association Conference, London, Ontario, Travel & Tourism Research Association. Retrieved from http://eresearch.qmu.ac.uk/2315

Reips, U. D. (2000). *The Web experiment method: Advantages, disadvantages, and solutions*. In M. H. Birnbaum (Eds.), *Psychological experiments on the Internet* (pp. 89-117). San Diego, CA: Academic Press.

Salkind, N. J. (2011). *Statistics for people who (think they) hate statistics* (*4th edition*). London: Sage.

Santos, J. R. A. (1999). Cronbach's alpha: A tool for assessing the reliability of scales. *Journal of Extension*, *37*(2), 1-5.

Schiffman, L.G. & Kanuk, L.L. (2000). *Consumer Behavior* (7th ed.). Wisconsin, US: Prentice Hall

Schauster, E. E., Ferrucci, P., & Neill, M. S. (2016). Native advertising is the new journalism: How deception affects social responsibility. *American Behavioral Scientist*, *60*(12), 1408-1424. doi: 10.1177/0002764216660135

Spears, N., & Singh, S. N. (2004). Measuring attitude toward the brand and purchase intentions. *Journal of Current Issues & Research in Advertising*, *26*(2), 53-66. doi: 10.1080/10641734.2004.10505164

Stapel, J. (1994). A brief observation about likability and interestingness of advertising. *Journal of Advertising Research, 34*(2), 79-80.

Statista. (2017). *Statista Survey Couponing 2016*. Statista.

Swant, M. (2016). *Twitter Says Users Now Trust Influencers Nearly as Much as Their Friends*. *Adweek.com*. Retrieved from http://www.adweek.com/digital/twitter-says-users-now-trust-influencers-nearly-much-their-friends-171367/

Toluna Quicksurveys. (2016). *TABS Beauty Study 2016*. TABS Analytics.

Tomkovick, C., Yelkur, R., & Christians, L. (2001). The USA's biggest marketing event keeps getting bigger: An in-depth look at Super Bowl advertising in the 1990s. *Journal of Marketing Communications, 7*(2), 89-108. doi: 10.1080/13527260121725

Trefis. (2017). Trefis.com.

Walker, D., & Dubitsky, T. M. (1994). Why liking matters. *Journal of Advertising Research*, *34*(3), 9-19.

Wei, M. L., Fischer, E., & Main, K. J. (2008). An examination of the effects of activating persuasion knowledge on consumer response to brands engaging in covert marketing. *Journal of Public Policy & Marketing, 27*(1), 34-44. doi: 10.1509/jppm.27.1.34

Wheeler, S. C., Briñol, P., & Hermann, A. D. (2007). Resistance to persuasion as self-regulation: Ego-depletion and its effects on attitude change processes. *Journal of Experimental Social Psychology*, *43*(1), 150-156. doi: 10.1016/j.jesp.2006.01.001

Williams, J. B., & Jacobs, J. S. (2004). Exploring the use of blogs as learning spaces in the higher education sector. *Australasian Journal of Educational Technology*, *20*(2), 232-247.

Wojdynski, B. W., & Evans, N. J. (2016). Going native: Effects of disclosure position and language on the recognition and evaluation of online native advertising. *Journal of Advertising, 45*(2), 157-168. doi: 10.1080/00913367.2015.1115380

Women's Wear Daily. (2017). *Beauty Inc.*. Women's Wear Daily.

Woods, S. (2016). *# Sponsored: The Emergence of Influencer Marketing.* University of Tennessee Honors Thesis Projects

Wright, K. B. (2005). Researching Internet-based populations: Advantages and disadvantages of online survey research, online questionnaire authoring software packages, and web survey services. *Journal of Computer-Mediated Communication*, *10*(3). doi: 10.1111/j.1083-6101.2005.tb00259.x

Wright, C. (2017). Are Beauty Bloggers More Influential than Traditional Industry Experts?. *Journal of Promotional Communications*, *5*(3), 303-322.

Wu, K. (2016). YouTube marketing: Legality of sponsorship and endorsements in advertising. *Journal of Law, Business & Ethics. 22*, 59.

Wu, M., Huang, Y., Li, R., Bortree, D. S., Yang, F., Xiao, A., & Wang, R. (2016). A tale of two sources in native advertising: Examining the effects of source credibility and priming on content, organizations, and media evaluations. *American Behavioral Scientist, 60*(12), 1492-1509. doi: 10.1177/0002764216660139

Appendix: Survey

Dear participant,

Thank you for clicking the link to this survey! My name is Nadia, and I am 21 years old. I am a Dutch student in the international Media & Business Master's program at the Erasmus University in Rotterdam. In order to obtain my degree, I have to write a thesis. I have chosen to do a thesis in which my research is based on the results of this survey. The following survey will take about 5-10 minutes to complete. The research consists of an online experiment in which we investigate beauty & make-up behaviors and opinions, especially in relation to YouTube. You will be asked a few questions about your behaviors and opinions with regards to certain make-up brands.

Please be aware that your participation is completely voluntarily, meaning that you can quit at any time during your participation. Furthermore, your personal information will be kept strictly confidential and the findings of this survey will be used solely for the purpose of this thesis. Hence, your anonymity is guaranteed at any time. If you have any questions or comments about the survey and/or research, please send an email to nadiathesiserasmus@gmail.com.

- o I understand the above and agree in participating in this research.
- o No, I do not wish to participate in this survey.
  (If no, go to end of survey)

1. With which gender do you identify the most?
   - o Male
   - o Female
   - o Other, namely …
     _____
   - o Prefer not to share

2. How old are you?
   _____

3. In which country do you currently reside?
   _____

4. Do you have an interest in make-up and/or beauty?
   - o Yes
   - o No

5. How often do you watch YouTube videos related to make-up and/or beauty?
   - o Multiple times a day
   - o Once a day
   - o Multiple times a week
   - o Once a week
   - o Less than once a week
   - o Less than once a month
   - o Never, I don't watch YouTube videos related to make-up and/or beauty

6. Please indicate which of the make-up brands below you have heard of. You can give multiple answers.
   - o Elf Cosmetics
   - o Rimmel London
   - o Maybelline New York
   - o L'Oreal Paris
   - o Essence
   - o Make-Up Academy

7. Please indicate how well you are acquainted with Maybelline based on the scales below.

Unfamiliar	O O O O O	Familiar
Inexperienced	O O O O O	Experienced
Not knowledgeable	O O O O O	Knowledgeable

8. Have you ever bought a make-up product from Maybelline?
   - o Yes
   - o No
   - o Don't remember
     (If yes, go to question 9. If no, go to question 10. If don't remember, go to question 10.)

9. What type of product from Maybelline was your last purchase? If there was more than once product purchased at the same time, please indicate the product that you used the most.
   - o Mascara
   - o Eyeliner or eyeshadow
   - o Primer, foundation or powder
   - o Lip product (liner, gloss, lipstick, etc.)
   - o Concealer
   - o Beauty tools
   - o Other, namely

     _____
   - o Don't remember

10. Please describe your overall feeling about Maybelline based on the scales below.

Unappealing	O O O O O O O	Appealing
Bad	O O O O O O O	Good
Unpleasant	O O O O O O O	Pleasant
Unfavorable	O O O O O O O	Favorable
Unlikable	O O O O O O O	Likable

11. How likely is it that you will buy a primer and/or foundation from Maybelline?
    - o Extremely unlikely
    - o Moderately unlikely
    - o Slightly unlikely
    - o Neither likely nor unlikely
    - o Slightly likely
    - o Moderately likely
    - o Extremely likely

12. In the next part of this survey, you will get to see the first two minutes of a YouTube video by YouTuber Basically Britt. Please watch the full video before you continue with the questions. If you do not see a video pop up immediately, please press the screen once and it should start. If you are unable to watch the video for some reason, please press the arrow to continue.
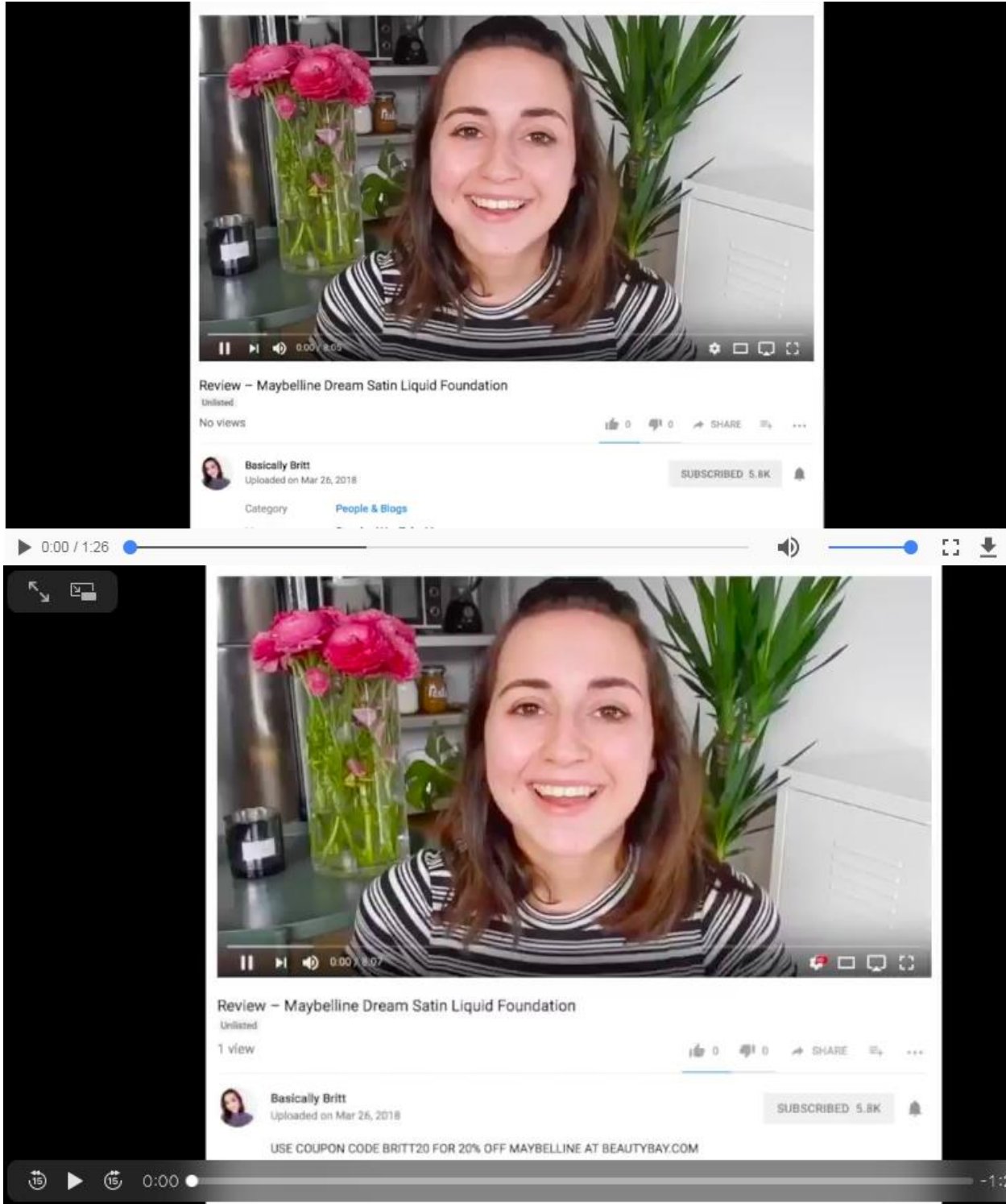
Review – Maybelline Dream Satin Liquid Foundation

Unlisted

1 view

👍 0   👎 0   ➤ SHARE   ☰₊   ...

Basically Britt
Uploaded on Mar 26, 2018

SUBSCRIBED 5.8K   🔔

THIS IS NOT A SPONSORED VIDEO

0:00

Review – Maybelline Dream Satin Liquid Foundation

Unlisted

2 views

👍 0   👎 0   ➤ SHARE   ☰₊   ...

Basically Britt
Uploaded on Mar 26, 2018

SUBSCRIBED 5.8K   🔔

THIS IS A SPONSORED VIDEO

0:00

Tutorial: How I Apply My Base | Maybelline Foundation
Unlisted
No views

👍 0   👎 0   ➡ SHARE

Basically Britt
Uploaded on Mar 26, 2018

SUBSCRIBED 5.8K

THIS IS A SPONSORED VIDEO

Tutorial: How I Apply My Base | Maybelline Foundation
Unlisted
6 views

👍 0   👎 0   ➡ SHARE

Basically Britt
Uploaded on Mar 26, 2018

SUBSCRIBED 5.8K

Category   People & Blogs

13. Were you able to watch the full 2 minutes of the video properly?
    o   Yes
    o   No, because

    _____
    (If yes, go to question 14. If no, go to end of survey)

14. What product was used in the video?
    o   A lipstick
    o   A foundation
    o   An eyeshadow palette

15. To what extent do you agree with the following statement:
    "This is the kind of video I tend to watch online."
    o   Strongly disagree
    o   Disagree
    o   Somewhat disagree
    o   Neither agree nor disagree
    o   Somewhat agree
    o   Agree
    o   Strongly agree

16. How much did you like the video?
    o   Dislike a great deal
    o   Dislike somewhat
    o   Neither like nor dislike
    o   Like somewhat
    o   Like a great deal

17. How likely are you to buy a primer and/or foundation from Maybelline?
    o   Extremely unlikely
    o   Moderately unlikely
    o   Slightly unlikely
    o   Neither likely nor unlikely
    o   Slightly likely
    o   Moderately likely
    o   Extremely likely

18. Please indication how well you think the following words describe the video in its entirety.

	Not well at all	Slightly well	Moderately well	Very well	Extremely well
Accurate	o	o	o	o	o
Authentic	o	o	o	o	o
Believable	o	o	o	o	o

19. Please indicate to what extent you agree/disagree with the statements below.
   a. *The YouTuber in the video is honest.*
o   Strongly disagree
o   Somewhat disagree
o   Neither agree nor disagree
o   Somewhat agree
o   Strongly agree

   b. *The YouTuber in the video is dependable.*
o   Strongly disagree
o   Somewhat disagree
o   Neither agree nor disagree
o   Somewhat agree
   Strongly agree

   c. *The YouTuber in the video is reliable.*
o   Strongly disagree
o   Somewhat disagree
o   Neither agree nor disagree
o   Somewhat agree
o   Strongly agree

   d. *The YouTuber in the video is sincere.*
o   Strongly disagree
o   Somewhat disagree
o   Neither agree nor disagree
o   Somewhat agree
   Strongly agree

   e. *The YouTuber in the video is trustworthy.*
o   Strongly disagree
o   Somewhat disagree
o   Neither agree nor disagree
o   Somewhat agree
o   Strongly agree

20. Please indicate to what extent you disagree/agree with the statements below.
   a. *The YouTuber in the video is knowledgeable.*
o   Strongly disagree
o   Somewhat disagree
o   Neither agree nor disagree
o   Somewhat agree
o   Strongly agree

b. *The YouTuber in the video is skillful.*
o Strongly disagree
o Somewhat disagree
o Neither agree nor disagree
o Somewhat agree
o Strongly agree

c. *The YouTuber in the video is experienced.*
o Strongly disagree
o Somewhat disagree
o Neither agree nor disagree
o Somewhat agree
o Strongly agree

d. *The YouTuber in the video is qualified.*
o Strongly disagree
o Somewhat disagree
o Neither agree nor disagree
o Somewhat agree
o Strongly agree

e. *The YouTuber in the video is an expert.*
o Strongly disagree
o Somewhat disagree
o Neither agree nor disagree
o Somewhat agree
o Strongly agree

21. Please describe your overall feeling of Maybelline based on the scales below.

Unappealing	O O O O O O O	Appealing
Bad	O O O O O O O	Good
Unpleasant	O O O O O O O	Pleasant
Unfavorable	O O O O O O O	Favorable
Unlikable	O O O O O O O	Likable

22. What do you think this survey aimed to research?
_____

23. Thank you for participating in this survey. If you wish to be informed about the results of this research, please send an email to nadiathesiserasmus@gmail.com