



LUND UNIVERSITY
School of Economics and Management

Are 5-Second Online Video Advertisements Effective?

Examining the relationship between advertisement length and
advertisement effectiveness

by

Jokūbas Augustas Kuprevičius and Anna Schuster

May 2017

Master's Programme in International Marketing and Brand
Management

Supervisor: Burak Tunca
Examiner: Johan Jansson

Abstract

5-second online video advertisements have emerged in the advertising scene as an ultra-short, theory defying phenomenon and scholars have yet to fully study its effectiveness. Up until recently, marketers have understood that advertisement effectiveness depends upon advertisement length, justifying the use of longer advertisement forms of 15-seconds or more. With the introduction of 5-second online video advertisements, one questions whether such a format is effective.

The effectiveness of 5-second online video advertisements was examined with respect to brand recall, message recall and advertisement likeability. More specifically, the effects of advertisement length, and interruption as a moderating variable were investigated for a 5-second online video advertisement. A 15-second version of the same advertisement was used for comparison reasons. An experiment was conducted using an online survey tool, with participants asked to view a YouTube video with an advertisement. Loglinear analysis and two-way analysis of variance were used to analysis results for each of the variables of advertisement effectiveness in isolation. Findings indicated that 5-second online video advertisements were relatively inadequate at producing brand recall, message recall or advertisement likeability in comparison to their longer versions. Interruption was found to moderate the relationship between advertisement length and effectiveness. However, results for brand recall, message recall and advertisement likeability remained weak for the 5-second version. Interruption had a much larger moderating effect for the 15-second version, suggesting that 5-second advertisements are hindered by their length. Resultantly, 5-second advertisements were deemed too short to be effective with respect to brand recall, message recall and likeability.

Keywords: Online video advertisements, 5-second, 15-second, effectiveness, advertisement length, interruption, brand recall, message recall, likeability, pre-roll, mid-roll, YouTube

Acknowledgements

First and foremost, we would like to express our thanks to our supervisor, Burak Tunca for his guidance and inspiration throughout the writing of this master thesis, for his patience in teaching us, and his valuable knowledge. His presence aided both our struggles and aspirations, and for that this thesis would not be the same.

Besides our supervisor, we would like to thank our professors who fuelled our curiosity, encouraging us to be critical, and passionate, about our work. In particular, we give our sincere thanks to Tony Marañon for his selfless support and kind spirit.

Last but not least, we would like to thank our families who provided us with the support to follow our dreams.



Jokūbas Augustas Kuprevičius



Anna Schuster

Table of Contents

- 1 Introduction 1**
- 1.1 Background 1
 - 1.1.1 Advertising 1
 - 1.1.2 Video Advertising 1
 - 1.1.3 Video Advertising on YouTube 2
 - 1.1.4 Video Advertising Terminology 2
- 1.2 Problem Discussion 3
- 1.3 Research Aim and Objectives 4
- 1.4 Research Questions 5
- 1.5 Outline of the Thesis 5
- 2 Theoretical Review 6**
- 2.1 A Conceptual Model of Advertising Effectiveness 6
- 2.2 Advertisement Length and Brand Recall 8
- 2.3 Advertisement Length and Message Recall 9
- 2.4 Advertisement Length and Advertisement Likeability 10
- 2.5 Interruption as a Moderating Variable 11
- 2.6 Overview of Theoretical Framework and Hypotheses 13
- 3 Methodology 14**
- 3.1 Research Approach 14
- 3.2 Research Design 14
- 3.3 Data Collection 15
 - 3.3.1 Sample 15
 - 3.3.2 Sampling Technique 16
 - 3.3.3 Collection Method 16
 - 3.3.4 Questionnaire 16
 - 3.3.5 Selection of Original YouTube Video 17
 - 3.3.6 Selection of Video Advertisement 18
 - 3.3.7 Pretest 18
 - 3.3.8 Procedure 19
 - 3.3.9 Measurements 21
- 3.4 Data Processing 22
 - 3.4.1 Assumptions of Loglinear Analysis 24
 - 3.4.2 Assumptions of Two-Way ANOVA 25

3.5	Validity and Reliability	26
3.6	Ethical Considerations.....	28
4	Results	29
4.1	Results for Brand Recall	29
4.1.1	Interpretation of Brand Recall Results	31
4.2	Results for Message Recall	33
4.2.1	Interpretation of Message Recall Results	35
4.3	Results for Advertisement Likeability	36
4.3.1	Interpretation of Advertisement Likeability Results	38
4.4	Overview of Tested Hypotheses	38
5	Discussion	40
5.1	The main effect of advertisement length on advertisement effectiveness	40
5.2	The interaction of advertisement length, interruption and advertisement effectiveness	42
6	Conclusion	44
6.1	Research Aims and Objectives.....	44
6.2	Theoretical and Practical Implications	44
6.2.1	Theoretical Implications.....	44
6.2.2	Practical Implications	45
6.3	Limitations and Future Research.....	45
7	References	47

List of Tables

Table 1. Experimental groups and assigned advertisements	20
Table 2. Survey questions and their purpose.....	21
Table 3. Categorisation of variables used in the experiment	22
Table 4. Overview of the sample demographics	23
Table 5. Overview of distribution of control variables	24
Table 6. Brand recall: Testing the assumptions of loglinear analysis.....	25
Table 7. Message recall: Testing the assumption of loglinear analysis	25
Table 8. Brand Recall: Cell counts and residuals.....	29
Table 9. Brand Recall: K-Way and higher-order effects	30
Table 10. Brand recall: Partial associations	30
Table 11. Brand recall: Parameter estimates	31
Table 12. Brand recall: Cell counts and residuals based on the adjusted model.....	31
Table 13. Brand recall: Cross-tabulation with frequencies	32
Table 14. Message recall: Cell counts and residuals.....	33
Table 15. Message recall: K-Way and higher-order effects.....	34
Table 16. Message recall: Partial associations	34
Table 17. Message recall: Parameter estimates.....	35
Table 18. Message recall: Cell counts and residuals based on the adjusted model	35
Table 19. Message recall: Cross-tabulation with frequencies.....	36
Table 20. Advertisement likeability: Descriptive statistics.....	37
Table 21. Advertisement likeability: Test of between-subjects effects	38
Table 22. Tested hypotheses and their status	39

List of Figures

- Figure 1. A conceptual framework..... 8
- Figure 2. A theoretical framework 13
- Figure 3. Screenshots of video advertisement and YouTube video used in the experiment.... 18
- Figure 4. Advertisement likeability: Testing the assumptions of two-way ANOVA 26
- Figure 5. Brand recall: Graphical presentation of frequencies..... 32
- Figure 6. Message recall: Graphical presentation of frequencies 35
- Figure 7. Advertisement likeability: Graphical presentation of means..... 38

1 Introduction

1.1 Background

1.1.1 Advertising

Advertising is one of the oldest and most recognizable forms of promotion, which often makes itself the face of the brand it promotes (McCarthy, 1960). Naturally, marketers flock to using this powerful tool, resulting in the advertising scene being restrictively competitive. Consequently, marketers and advertisers alike are continually inventing new ways to retain advertisement recognition in this cluttered and oversaturated market (Jeong, 2015; Newstead & Romaniuk, 2009).

A separate challenge marketers are facing is the ever-changing environment of advertising, causing the term itself to evolve. After studying these changes, Richards and Curran (2002) have developed a definition that includes advertising's fundamental elements. According to them, "advertising is a paid, mediated form of communication from an identifiable source, designed to persuade the receiver to take some action, now or in the future" (Richards & Curran, 2002, p.74). Perhaps the most notable change in advertising is visible through the variety of channels it spanned in the past compared to the amount it spans today. Until recently, advertisements were displayed mostly through print media, radio and television (Singh & Cole, 1993). Advertising through the Internet has become one of the most popular channels of promotion chosen by brands today. In 2016, online advertising accounted for 33,8% of advertising spending worldwide, just barely behind television 35,6% (Statista, 2017a). Unsurprisingly, online advertising is estimated to outrun television advertising in the upcoming few years (Statista, 2017a). Notwithstanding, video advertisements have proven to be a competitive category of advertising within all channels, both offline and online. With it being the main type of advertising used in television, and picking up speed online, video advertising is in especially high demand - thus, making it of great importance for advertisers to keep up with its changes.

1.1.2 Video Advertising

One of the most significant changes that video advertisements have undergone is their shift in length. This change was first spotted within the offline medium since online advertising had not emerged before the birth of commercial Internet usage in the 1990s (Li & Lo, 2015). In the 1950's, the early days of television, advertisement length was standardised at 60 seconds (Jeong, 2015). However, in the 1960s, length was experimented with, resulting in a new 30-second advertisement that quickly became the new standard for television advertisements (Jeong, 2015). In the beginning of the 1980s, a U.S. television network introduced its shortest advertisement yet, of 15 seconds, which quickly gained popularity worldwide (Singh, Linville & Sukhdial, 1995). Since then, shorter alternatives have continually grown in popularity (Li & Lo, 2015). With the introduction of online video advertising, it did not come as a surprise that shorter video formats were going to be used on the Internet as well. Today, it is not uncommon to encounter online video advertisements that are only 10 seconds or less. This

trend has been accompanied by decreasing attention spans. A recent study by Microsoft (2015) found that average attention spans fell to eight seconds in 2015, a decrease of four seconds from 2013. Together with the fact that the advertising scene is becoming more saturated, the length of online video advertisements is said to decrease now and in the future.

1.1.3 Video Advertising on YouTube

When it comes to online video advertising, there is great variety in both length and place. A reason behind this trend is that the Internet is significantly more fragmented than television, meaning that there is not one entity that controls advertising, but many. As such, each website may decide what kind of video advertisements it displays. Some examples of video advertisement types include separate pop-up video advertisements, or advertisements placed within another video, such as those on YouTube.

YouTube is the world's second largest search engine after Google and is the most popular video sharing website worldwide (Dehghani, Niaki, Ramezani & Sali, 2016; Gupta, Singh & Sinha, 2016). Millions of users are exposed to video advertisements on the website every day. Thus, YouTube is an important, if not the most important, channel to consider when analysing online video advertising. Today, marketers can choose from six different formats when placing advertisements on YouTube (YouTube, 2017a). Out of these six, two are particularly interesting because of the unprecedentedly short length they offer, named Bumper and Skippable Video advertisements (YouTube, 2017a). Bumper advertisements are non-skippable advertisements that are displayed before the original YouTube video commences and are up to 6 seconds in length (YouTube, 2017a). Skippable Video advertisements can be placed before, or in the middle, of the original YouTube video and can be longer in length. However, with Skippable Video advertisements, the viewer has the possibility to commence the YouTube video after 5 seconds (YouTube, 2017a). Resultantly, advertisers are increasingly producing advertisements that span no more than 5 seconds. Therefore, YouTube, the channel itself, is already implying that advertisers should make advertisements significantly shorter than the ones shown on television. One could potentially raise questions whether such significant reductions in advertisement length is needed and whether these new advertising formats are effective.

1.1.4 Video Advertising Terminology

In this paper, the terms 'video advertising' and 'video advertisements' are used when talking about audio-visual advertising, that is, content that includes both visuals and sound (Li & Lo, 2015). Such advertisements are often placed within video channels, such as YouTube, given the matching format. It is important to note that advertisements that do not contain sound or motion, such as moving banner ads, do not qualify as video advertisements.

Video advertisements are split into multiple categories concerning different aspects. One of these aspects is the type of medium the video advertisement is shown through. Therefore, the terms 'offline video advertising' and 'offline video advertisements' are used when referring to video advertisements shown through offline mediums such as television. 'Online video advertising', together with 'online video advertisements', are used for video advertisements shown through online channels such as YouTube or other online platforms.

To reduce confusion between terms, it should be mentioned that the paper refers to two different types of videos - the video advertisement and the original YouTube video.

Additionally, video advertisements are separated into two categories when talking about their placements with respect to the YouTube video. These are ‘pre-roll’ and ‘mid-roll’ advertisements. ‘Pre-roll’ advertisements refer to video advertisements that are shown before the video being viewed, and ‘mid-roll’ advertisements refer to interrupting video advertisements shown during the video (Li & Lo, 2015).

1.2 Problem Discussion

Video advertising as a marketing tool is well recognised in previous literature, and research into its effectiveness has an extensive history (Biel & Bridgwater, 1990; Fabian, 1986; Greene, 1991; Mord & Gilson, 1985; Newell & Henderson, 1998; Singh & Cole, 1993). Only recently have studies of video advertising directly addressed online channels (Goodrich, Schiller & Galletta, 2015; Li & Lo, 2015).

Different theories exist in the literature regarding the relationship between advertisement length and the effectiveness of advertising, and many circulate around three primary variables of effectiveness - brand recall (Martin-Santana, Reinares-Lara & Reinares-Lara, 2016; McCoy, Everard, Polak & Galletta, 2008), message recall (McCoy et al., 2008; Newstead & Romaniuk, 2009) and advertisement likeability (Smit, van Meurs & Neijens, 2006). It is now established from a variety of studies that all three variables are important in contributing to advertisement effectiveness. Together, these studies specifically observed video advertising in offline environments, such as television commercials, while no such scholars have tested the variables of brand recall, message recall and advertisement likeability in the online environment. Instead, authors have focused on studying brand recall, message recall and advertisement likeability of online advertisements in a fragmented manner (Goodrich, Schiller & Galletta, 2015; Li & Lo, 2015; Yoon, Choi & Sujin, 2011). With the introduction of YouTube, video advertising has migrated to online channels, providing the opportunity for new insights into how the effectiveness of video advertising will change as a result.

Moorman, Neijens and Smit (2002) reported that a significant difference exists between pre-roll and mid-roll advertisements in stimulating brand recall. Based on this finding, Li and Lo (2015) conducted two experiments that tested the effect of length on brand recall and interruption-based (mid-roll) advertisements on brand recall. However, the authors did not consider that interruption could be a moderating variable in altering the relationship between advertisement length and brand recall. Additionally, the authors did not test whether the two could have significant effects on other important variables such as message recall and advertisement likeability. Prior to the work of Li and Lo (2015), the role of online interruption-based video advertising on brand recall was largely unknown.

Although online interruption-based video advertising remains quite a new research topic, the use of interruption in other forms of online advertising has attracted considerable research. Online interruption-based advertising has been implemented as a marketing strategy for many brands, usually as pop-up advertisements (non-video). Jankowski, Kazienko, Wątróbski, Lewandowska, Ziemia and Ziolo (2016) delved into how user experience is negatively affected given the oversaturation of pop-up advertisements. Online advertising has made interruption-based advertising a rather easy feat, and enables brands to be present and recognised virtually anywhere. Having that said, interruption-based advertising, both non-

video and video based, has become so popular that user experience has suffered as a result. This assumption is well presented by Heffernan (2009) who wrote that advertisements in the online environment are often seen as an imposition, causing indiscriminate negative emotions towards advertisements in general. One of the main limitations of current research on the topic of online video advertising is that researchers lack understanding of how this intrusive environment alters the similarity between offline and online advertisement effectiveness, particularly with respect to advertisement likeability.

Understanding the relationship between advertising length and effectiveness is an essential element to the success of an advertisement and a brand's profitability (Newstead & Romaniuk, 2009). Additionally, online advertising is becoming the dominant form of advertising in the 21st century (Chan, Jiang & Tan, 2010). Specifically, interruption-based online advertising is said to be the most widely used advertising technique, regardless of the channel (Chan, Jiang & Tan, 2010), and is why further study on the topic is essential. The literature on online advertising has highlighted that advertisements are becoming shorter (Newstead & Romaniuk, 2009) but there lacks understanding of whether advertisements are becoming less effective as a result.

In summary, advertising effectiveness has said to positively correlate with advertisement length. The longer the advertisement, the more effective it becomes (Li & Lo, 2015; Newstead & Romaniuk, 2009). With the now oversaturation of the advertising scene, as well as the counter-force of shorter attention spans, advertisements have become shorter than ever (Goodrich, Schiller & Galletta, 2015; Li & Lo, 2015). This trend has been felt in various channels, especially on the Internet, given that it is the fastest growing medium for advertisers of the 21st century (Chan, Jiang & Tan, 2010; Statista, 2017a). The introduction of the Internet has meant that the reach of advertising has extended to every second of our lives, but in its wake produced many challenges. For one, the ability to generate brand recall has become a challenge as consumers process more advertisements per day than ever before (Li & Lo, 2015). Moreover, marketers have attempted to increase message recall and advertisement likeability but have been hindered by the limited attention spans of consumers as well as cost constraints (Microsoft, 2015; Newstead & Romaniuk, 2009). Therefore, marketers have turned to shorter interruption-based advertisements as tools to increase attention and effectiveness. In light of this, a new phenomenon has occurred in which brands are utilising 5-second advertisements, seeming to go against the above mentioned findings regarding the superiority of longer advertisements. Similarly, there lacks substantial research in understanding the relationship between advertisement length and effectiveness in the online medium.

1.3 Research Aim and Objectives

The objective of this study is to investigate the phenomenon of 5-second online video advertisements. This study aims to shed light on their effectiveness by measuring how 5-second advertisements compare to 15-second advertisements, another popular format that has been studied and used in the advertising industry (Singh, Linville & Sukhdial, 1995). Additionally, this paper will explore whether interruption moderates the relationship between advertisement length and effectiveness. The specific objective of this is to explore environments where effectiveness may be enhanced. The results of the study will be

compared to previous research, with the aim of contributing to the understanding of the effectiveness of ultra-short online video advertisements.

As online advertising becomes the dominant type of promotion in the 21st century (Statista, 2017a), understanding its effectiveness is imperative to the success of online video advertising in general. The methods used in this paper also contribute to understanding whether online channels changes the relationship between variables of video advertising effectiveness. The knowledge that this study aims to reveal could influence marketers to rethink their traditional models and provide more effective campaigns given the new environment.

1.4 Research Questions

RQ1: What is the relative effectiveness of 5-second advertisements compared to 15-second advertisements?

RQ2: What is the influence of interruption on the relationship between advertisement length and effectiveness?

1.5 Outline of the Thesis

To test the relationship between advertisement length and effectiveness, the following chapter provides theories of which hypotheses will follow. The primary variables testing advertisement effectiveness will be introduced, and the moderating variable of interruption will be elaborated upon. Following the theoretical framework will be the methodology in which the experimental study will be presented, as well as techniques of data collection, data processing and limitations. The corresponding results section will present the findings of the experiment according to the hypotheses tested. Consequently, an analysis will follow in which existing theory will be drawn upon to aid discussion of the results. A conclusion will follow, in which the study is summarised. The conclusion will also include limitations of the study and implications for further research.

2 Theoretical Review

2.1 A Conceptual Model of Advertising Effectiveness

While sales are often seen as the most important metric in evaluating advertising effectiveness, researchers tend to use more specific variables such as brand recall, message recall and advertisement likeability when measuring effectiveness (Newstead & Romaniuk, 2009). In fact, each of these variables provides unique contributions to the understanding of what makes advertisements effective, and is why they are often measured individually (Haley & Baldinger, 2000).

Advertising theory has its roots in psychology with the aim of understanding how components of an advertisement affect people (Laczniak, 2015). The discipline itself is more practical than theoretical in its nature, and as a result, research on the topic is ever changing. Paired with the digital influence that the Internet has bestowed, advertising is undergoing significant changes that extend previous knowledge on the topic. However, at the very roots of advertising theory are concepts taken from disciplines such as psychology, which provide insights into new phenomena of advertising we see today, such as ultra-short 5-second video advertisements (Laczniak, 2015).

One such model that is indeed based on psychology theory is AIDA (Strong, 1925). The model developed by Strong (1925) is constructed of (A) attention, (I) interest, (D) desire and (A) action. Of particular interest to the study are the components of attention and interest, for which, when measured, provide critical insights into whether the advertisement has been successfully received by the viewer (Strong, 1925). Accordingly, this study defines the effectiveness of an advertisement by its success in obtaining the attention and interest of the viewer. This definition is similar to that of Gupta, Singh and Sinha (2016) who studied the effectiveness of advertisements according to both attention and interest.

Attention is often measured through brand and message recall and is based on the psychological foundation of learning (Gupta, Singh & Sinha, 2016). The theory related to these variables is named 'total time hypothesis', and it suggests that a linear relationship exists between the amount of exposure to the advertisement and its recall (Pechmann & Stewart, 1988; Rethans, Swasy & Marks, 1986). The theory is grounded by the knowledge that longer advertisements provide repetition of the brand name and message, which aids learning (Fabian, 1986; Mord & Gilson, 1985; Singh & Cole, 1993). On a similar note, the encoding of information (regarding brand and message recall), is facilitated by this repeated information (Singh & Cole, 1993). These assumptions have been validated in previous research, with findings suggesting the superiority of longer advertisements in terms of brand and message recall (Newstead & Romaniuk, 2009).

The second component of the AIDA model, labelled as interest, provides a foundation of knowledge for the third variable in question - advertisement likeability. Gupta, Singh and Sinha (2016) define interest by whether or not the advertisement is deemed likeable. Characteristics of an advertisement that influence likeability include music, storytelling,

visual effects and colours (Gupta, Singh & Sinha, 2016). Likeability is an important variable to consider because of its influence on learning, especially in respect to memory (Albers-Miller & Stafford, 1999). As a result, the two components of the AIDA model are intertwined and equally important (Gupta, Singh & Sinha, 2016). In fact, Gupta, Singh and Sinha (2016) make a point of stating that “every advertisement that has been able to catch attention may not be ... liked by the consumer”. It is the combination of both attention and interest that stimulates the corresponding components of desire and action. Such is why it is necessary to consider both attention and interest when measuring advertisement effectiveness.

An equally important theory relating to both attention and interest is the ‘message learning theory’ by Hovland (1953). The theory is useful in understanding the effects of advertising with respect to attention and interest and is, in fact, the focal point of the theory. Accordingly, persuasion is obtained when people are passed through a process of attention, comprehension, yielding and retention of the message (Hovland, 1953). Of interest to this study is the assumption that this persuasion depends on learning. As such, the more people learn, the more they are persuaded. Applying this to the study, it is assumed that longer advertisements, with longer exposure, are more persuasive than shorter advertisements.

In summary, traditional advertising theories suggest that advertisement length and advertisement effectiveness are positively correlated (Albers-Miller & Stafford, 1999; Hovland, 1953; Pechmann & Stewart, 1988; Rethans, Swasy & Marks, 1986; Strong, 1925). However, few studies have examined how interruption alters the relationship between the two variables within online video advertising. It is known that interruption plays a role in disrupting the effectiveness of non-video advertising (Chan, Jiang & Tan, 2010; Jankowski et al., 2016), but in such cases, length was not a relevant variable. Understanding the role of interruption in the online video environment can lead to new insights into what makes online video advertising effective, and if traditional notions remain relevant.

Assumptions relating to interruption-based advertisements are based on ‘goal pursuit’ theory by Laran, Janiszewski and Salerno (2016) as well as the ‘elaboration likelihood model of Persuasion’ (ELM) by Petty, Cacioppo and Schumann (1983). Goal pursuit concerns actively processing information on the quest for relevance (Laran, Janiszewski & Salerno, 2016). As such, people who are goal-oriented are more prone to learning (Ha & McCann, 2008). In the environment of YouTube, goal-orientation comes in the form of watching informative videos which involves actively processing information (Li, Edwards & Lee, 2002). When disrupted, viewers adopt a similar level of cognitive activity for the advertisement as the original video (Li, Edwards & Lee, 2002). As such, learning of the advertisement content is heightened when the advertisement is interruption-based because of this spill-over effect (Li, Edwards & Lee, 2002). This superior mental engagement provides the assumption that interruption positively moderates the relationship between advertisement length and recall (Li, Edwards & Lee, 2002). However, an important distinction should be made between brand/message recall and likeability. The latter concerns one's attitude towards the advertisement, which, intuitively, does not perform well in interruption-based environments. Negative psychological reactions brought on by interruption, such as the feeling of intrusiveness, annoyance and constriction, weaken the general likeability of an advertisement (Goodrich, Schiller & Galletta, 2015). Petty, Cacioppo and Schumann's (1983) ELM model provides some further explanation on this effect. The authors suggest that a change in attitude depends on the devotion of cognitive effort towards the message, or in this case the advertisement. When a person is particularly cognitively active, referred to by Petty, Cacioppo and Schumann (1983) as taking the ‘central route’, people processes messages more critically. Such is the case when

processing an informative online video (Li, Edwards & Lee, 2002). Paired with the fact that interruption-based advertisements form negative psychological reactions, people are more prone to criticism (Li, Edwards & Lee, 2002). In this respect, interruption is assumed to invert the relationship between length and likeability.

Based on the conceptual model in Figure 1, we now discuss the research surrounding the length of a video advertisement and its influence on the variables of brand recall, message recall and advertisement likeability. In addition, we examine the role of interruption as a moderating variable.

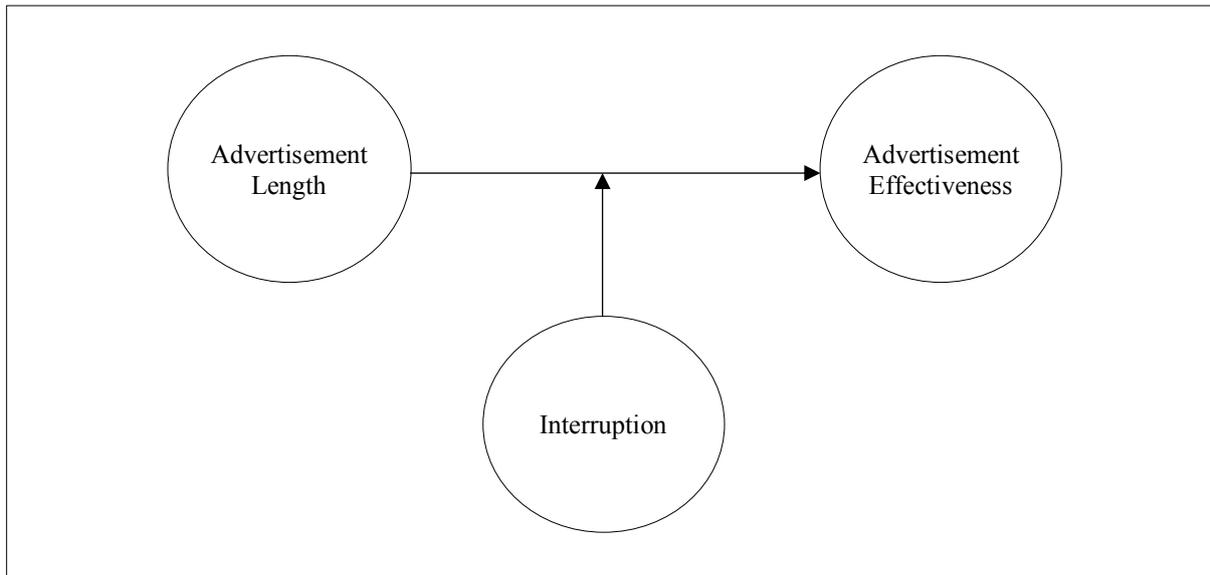


Figure 1. A conceptual framework

2.2 Advertisement Length and Brand Recall

An important objective of advertising, or any marketing effort for that matter, is brand recall (Kapferer, 1997; Keller, 2008). Recall refers to the correct identification of a brand name when prompted by the product category (Kapferer, 1997; Keller, 2008). Brand recall stimulates future brand recognition and consumption, aiding the profitability of the brand (Newstead & Romaniuk, 2009). It is paramount to the success of any brand and is often considered the most important metric in advertising (Newstead & Romaniuk, 2009).

Brand recall concerns the component of attention in the AIDA model (Gupta, Singh & Sinha, 2016), and marketers have put focus on this particular component because it is essential to the success of the other components (Gupta, Singh & Sinha, 2016). Attention is dependent on advertisement length, as suggested by the 'total time hypothesis' (Pechmann & Stewart, 1988; Rethans, Swasy & Marks, 1986). The specific theory provides the assumption that the longer the advertisement, the more attention it receives (Pechmann & Stewart, 1988; Rethans, Swasy & Marks, 1986). As the message learning theory suggests, retention of information is a direct function of attention (Hovland, 1953). As such, longer advertisements provide the basis for increased attention, learning and, in turn, brand recall.

Previous research on the topic of advertisement length and brand recall has validated this theory. Within the scope of offline advertising, such as television, it was deemed the longer the advertisement the more effective the brand recall (Elsen, Pieters & Wedel, 2016). Arguments towards such a statement suggests that longer advertisements enable longer exposure time, which results in strengthened memory of the brand name (Danaher & Mullarkey, 2003; Goldstein, McAfee & Suri, 2011). Longer advertisements also allow for increased mentions of the brand name (Newstead & Romaniuk, 2009). In radio advertising, Allan (2007) studied the effect of advertisement length and brand recall, with findings supporting the familiar positive correlation between the two variables. In a similar study, McCoy et al. (2008) demonstrated that longer pop-up advertisements lead to increased brand recall. Additionally, longer advertisements contain persuasive techniques such as brand storytelling that strengthen the memorability of the advertisement and the personality of the brand (Lundqvist, Liljander, Gummerus & van Riel, 2012).

Based on the above findings, advertisement length and brand recall are positively correlated, leading to the following hypothesis:

H1a: 15-second online video advertisements generate better brand recall than 5-second online video advertisements.

2.3 Advertisement Length and Message Recall

Message recall is another objective of advertisers and marketers alike. It measures the effectiveness of communication within the advertisement, with focus on the general tagline or slogan of the brand (Newstead & Romaniuk, 2009). Several studies have shown that advertisement length and message recall are positively correlated (Fabian, 1986; Mord & Gilson, 1985; Singh & Cole, 1993).

The theory concerning message recall is identical to that used for brand recall and will be mentioned rather reservedly here. Instead, findings of previous research will be highlighted. As with brand recall, previous research on the topic of advertisement length and message recall has validated the theory. Message recall has been found to be dependent on advertisement length, with longer advertisements generating increased message recall. Newstead and Romaniuk (2009) found this relationship when specifically testing 15- and 30-second advertisements. The authors established that 15-second advertisements were 80% as effective as 30-second advertisements in this regard. Singh, Linville and Sukhdial (1995) produced a similar study, concluding that length plays an important role in retaining the advertisement claim. Newell and Henderson (1998) extended this study by using additional advertisement lengths, with findings confirming the strong relationship between the two variables of length and message recall.

Arguments towards this assumption suggest that shorter advertisements lack the repetition required for a memorable message (Singh & Cole, 1993). Singh and Cole (1993, p.92) state that “repetition facilitates better encoding”. Although repetition is not confined to longer advertisements, it certainly provides space to do so. As such, message recall is heightened by repetition, which is correlated with advertisement length. Another argument suggests that

learning depends on the amount of attention given to process information (Hovland, 1953). This would suggest that longer advertisements are superior to their shorter counterparts.

Based on the above findings, advertising length and message recall are positively correlated, leading to the following hypothesis:

H1b: 15-second online video advertisements generate better message recall than 5-second online video advertisements.

2.4 Advertisement Length and Advertisement Likeability

Advertisement likeability is a third metric used to measure advertising effectiveness. It refers to a positive emotional response to an advertisement (Newstead & Romaniuk, 2009). Data from several studies suggest that likeability is a valid variable to measure advertisement effectiveness (Greene, 1991; Mord & Gilson, 1985; Smit, van Meurs & Neijens, 2006). It contributes to the effectiveness of an advertisement because it enhances memorability, both of the brand name and the advertisement content (Gupta, Singh & Sinha, 2016). Arguments towards such a statement suggest that consumers are more attentive towards advertisements that they like, as validated by previous research (Biel & Bridgwater, 1990; Gupta, Singh & Sinha, 2016). The variable itself was said to be linked to campaign success, and if weak, may lead to its failure (Haley & Baldinger, 2000). As a result, likeability has been used as an important variable in measuring advertising effectiveness.

Within the scope of offline advertising, it was deemed the longer the advertisement, the more likeable it becomes. Mord and Gilson (1985) investigated the differential impact of length to likeability and found that longer television commercials were superior to shorter commercials in likeability scores. This finding explicitly assumes that shorter video advertisements lack interesting content, are less informative and more irritating than longer commercials (Gupta, Singh & Sinha, 2016; Mord & Gilson, 1985).

Additionally, researchers investigating advertisement length and effectiveness have compared 15-Second and 30-Second television commercials (Fabian, 1986; Mord & Gilson, 1985; Newstead & Romaniuk, 2009; Singh & Cole, 1993), given the limited variation of television commercials. Mord and Gilson (1985) found that shorter television advertisements enhances the feeling of clutter and reduce advertising effectiveness as a result. The study showed that viewers preferred longer advertisements while finding shorter advertisements more irritating (Mord & Gilson, 1985). However, a number of studies have also postulated that advertisement length and effectiveness are not strongly correlated. Fabian (1986) found that 15-second advertisements, while not as effective as their 30-second counterparts, still remain relatively on par with their counterparts in terms of recall, communication, and persuasion ability. Although half the length, 15-second advertisements are 75-85% percent as 30-second advertisements effective in respect to the aforementioned variables (Fabian, 1986). Singh and Cole (1993) theorise that a strong correlation ceases to exist because of moderating variables affecting the relationship between the two variables. Newstead and Romaniuk (2009) conducted an experiment that tested brand recall, message recall and likeability for 15-second versus 30-second advertisements. In their analysis, (Newstead & Romaniuk, 2009) concluded

that 30-second advertisements produced significantly higher likeability scores than 15-second advertisements shown on television.

Based on the above findings, advertising length and likeability are positively correlated, suggesting that longer advertisements are superior to shorter advertisements in terms of likeability. This leads to the following hypothesis:

H1c: 15-second online video advertisements generate better advertisement likeability than 5-second online video advertisements.

2.5 Interruption as a Moderating Variable

As previously stated, research on interruption in video advertising has provided some insights into how the relationship between advertisement length and advertisement effectiveness may change. Li and Lo (2015) found that interruption enhances the positive relationship between advertisement length and both brand and message recall. Li and Lo's (2015) findings are based on the assumption that watching videos online requires high levels of attention, and is validated by the goal-pursuit theory by Laran, Janiszewski and Salerno (2016). In this environment, consumers are cognitively active, much more so than in offline environments such as while watching television (Ha & McCann, 2008). Using platforms such as YouTube requires consumers to actively watch their computer screens, and thus involves active information processing (Ha & McCann, 2008). In addition, Chatterjee (2008) conducted an in-depth study on how mid-roll, interruption-based advertisements cannot be easily cognitively avoided.

It is well established in previous knowledge that interruption-based advertising has a general positive impact on the relationship between length and recall. Li and Lo (2015) claim that mid-roll advertisements attract significantly more cognitive attention than their pre-roll counterparts. The authors state that “viewers are expected to be more mentally engaged in the middle of the video than before or after it” (Li & Lo, 2015, p.209). This is because the attention given to the online video is spilled over when viewing the interruptive advertisement (Lloyd & Clancy, 1991). Krugman (1983) supports Lloyd and Clancy's (1991) findings, claiming that this relationship exists also in offline environments, specifically in television. Additionally, Moorman, Neijens and Smit (2002) observed that interruption-based advertising stimulated better recall, both in terms of brand name and advertisement message. In summary, mid-roll online video advertisements attract higher levels of attention, allowing for increased advertising effectiveness.

Based on the above assumptions, interruption positively affects the relationship between advertising length and the variables of brand and message recall, due to stimulated cognitive attention. This leads to the development of the following hypotheses:

H2a: Interruption strengthens the positive relationship between online video advertisement length and brand recall.

H2b: Interruption strengthens the positive relationship between online video advertisement length and message recall.

Interruption, however, is not expected to produce only positive effects. Negative psychological reactions, such as the feeling of intrusiveness, annoyance and constriction weaken the general likeability of an advertisement (Goodrich, Schiller & Galletta, 2015). This supports intuitive arguments that suggest our distaste for interrupting advertisements (Heffernan, 2009). Theories suggesting that interruption inverts the relationship between length and likeability are generated by Petty, Cacioppo and Schumann's (1983) ELM model and provides some further explanation on this effect. The authors suggest that a change in attitude depends on the devotion of cognitive effort towards the message, or in this case the advertisement. When a person is particularly cognitively active, referred to by Petty, Cacioppo and Schumann (1983) as taking the 'central route', people process messages more critically. Such is the case when processing an informative online video (Li, Edwards & Lee, 2002). Paired with the fact that interruption-based advertisements form negative psychological reactions, people tend to be more critical (Li, Edwards & Lee, 2002). The annoyance effect is said to enhance in channels where consumers are more cognitively involved, such as when viewing audio-visuals (Speck & Elliott, 1997). Given that online users are much more goal oriented than offline users (Ha & McCann, 2008), interruption-based online advertisements are seen to be more intruding, annoying and as a result, more unlikeable than in offline mediums (Li, Edwards & Lee, 2002). The inversion of the positive relationship between advertisement length and advertisement likeability occurs because shorter advertisements weaken the negative implications that interruption bestow (McCoy et al., 2008). Thus, interruption can have direct negative effects on the likeability of online video advertisements.

Based on the above assumptions, interruption is said to invert the relationship between advertising length and likeability, due to the irritation and annoyance caused by the interruption. This leads to the development of the final hypothesis that suggests that longer, 15-second advertisements will produce lower likeability scores than 5-second advertisements.

***H2c:** Interruption inverts the positive relationship between online video advertisement length and advertisement likeability.*

2.6 Overview of Theoretical Framework and Hypotheses

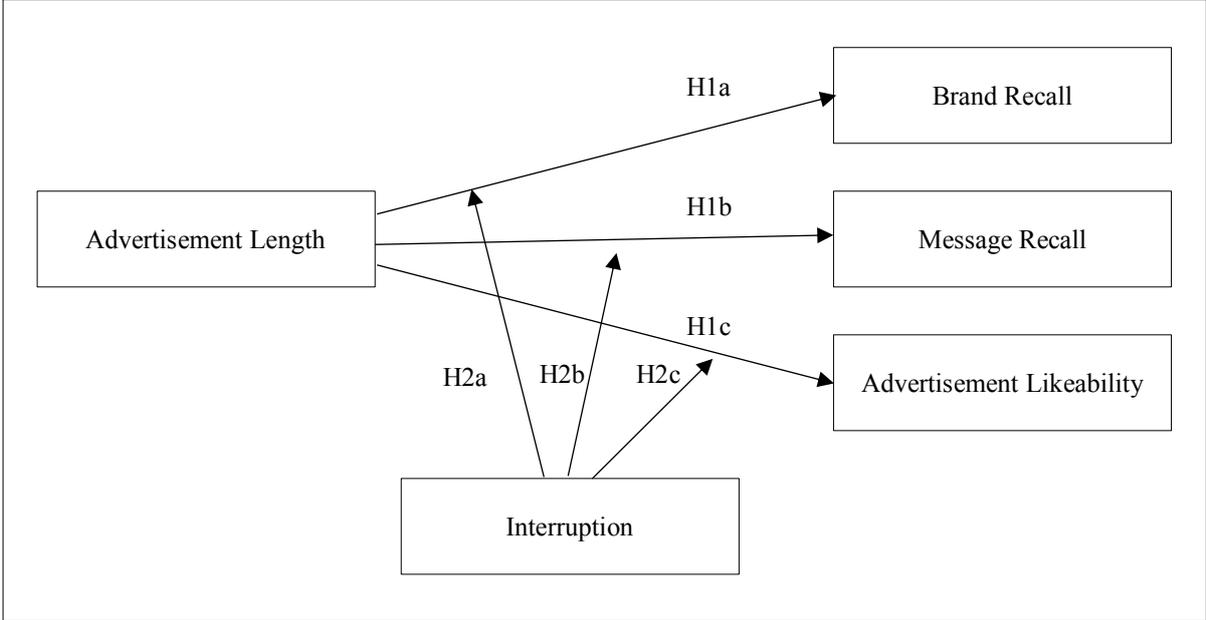


Figure 2. A theoretical framework

We have developed several hypotheses relating to the relationship between advertising length and brand recall, message recall and advertisement likeability. Our hypotheses are directed by the theoretical framework in Figure 2. More specifically, we transfer the theory of offline video advertising to test whether the typical relationship between advertisement length and advertisement effectiveness remains in the online environment. We hypothesise that the relationship between advertisement length and effectiveness remains similar for all variables. However, when it comes to interruption as a moderating variable, the relationship between advertisement length and advertisement likeability becomes inverted, causing shorter advertisements to become more effective than their longer counterparts. On a similar note, it is hypothesised that interruption will strengthen the positive relationship between advertisement length and the recall variables.

3 Methodology

3.1 Research Approach

“I didn’t think; I experimented” - Wilhelm Roentgen

The consideration of one’s philosophical stance is of great importance when conducting research and very much guides the chosen research method (Easterby-Smith, Thorpe & Jackson, 2015). Of the two philosophical paradigms of social constructionism and positivism, the approach taken by us is positivism. Positivist ideologies function on the idea that reality can be measured through objective rather than subjective methods (Easterby-Smith, Thorpe & Jackson, 2015). A positivistic approach provides several implications to the research design, of which this paper adopts. Firstly, explanations must generally demonstrate causality (Easterby-Smith, Thorpe & Jackson, 2015). Given this, the research aim - to understand the relationship between advertisement length, interruption and the variables of advertisement effectiveness - is fitting towards our positivistic philosophy. Secondly, research generally progresses through hypotheses and deductions of which are tested using large numbers of randomly selected subjects (Easterby-Smith, Thorpe & Jackson, 2015). The outcome of such an approach is the confirmation of theory (Easterby-Smith, Thorpe & Jackson, 2015).

As with most social researchers, philosophical stances and research designs are often inherited from previously published researchers of the topic (Easterby-Smith, Thorpe & Jackson, 2015). There have been several studies devoted to understanding the relationship between advertisement length and advertisement effectiveness for which our study is influenced (Goodrich, Schiller & Galletta, 2015; Li & Lo, 2015; Newstead & Romaniuk, 2009). A quantitative, experimental design is one of the most common procedures for determining the effectiveness of advertisements and has shown little change between years of research. This method is particularly useful in studying the relationship between variables and is complemented by the deductive approach of the paper. A deductive research approach often accompanies quantitative research methods given its aspiration towards generalisability (Ghauri & Grønhaug, 2005; Saunders, Lewis & Thornhill, 2007). The superiority of such an approach is that it enables the development of an objective view of social reality and is grounded by existing theory (Bryman & Bell, 2011).

3.2 Research Design

A research design is a framework for collecting and analysing data for a research project (Burns & Burns, 2008; Malhotra, 2010). It is the basis for the decisions and methods used throughout the study and is generally classified into two separate categories, those being exploratory and conclusive designs. Both of which are heavily influenced by the nature of the research question (Malhotra, 2010). For more loosely defined research questions, an

exploratory design is necessary (Malhotra, 2010). However, for more formal and structured research, a conclusive design would be apt (Malhotra, 2010).

Accordingly, our research design is influenced by the nature of quantitative research that calls for conclusive designs in the form of descriptive or causal research (Ghauri & Grønhaug, 2005). Given the research problem, the study adopts a conclusive research design. Of the conclusive research designs, the study adopts the principles of causal research with the aim to substantiate a cause-and-effect relationship (Malhotra, 2010). It is the objective of the paper to understand how the dependent variables (brand recall, message recall and advertisement likeability) respond to the changes in the independent variable (advertisement length). Additionally, the influence of the moderating variable (interruption) on this relationship will be tested. The independent and moderating variables will be manipulated in a controlled, experimental environment in order to test the resulting causal relationship (Malhotra, 2010). It should be noted that the authors respect the scientific meaning of causation, whereby “X [independent variable] is only one of a number of possible causes of Y [dependent variable]” (Malhotra, 2010, p.218). Saying that, the conclusions made from the experiment will by no means be concrete. In fact, as Malhotra (2010, p.218) states: “we can never prove that X is a cause of Y. At best, we can infer that X is a cause of Y”. Such is the aim of this study.

3.3 Data Collection

The method of data collection is strongly influenced by many factors, of which include (1) the nature of the population at hand, such as age and geographical dispersal, (2) the length of the experiment, as well as (3) financial and other resources, such as time and the number of interviewers available. The main aim of the data collection method is to select an approach that attracts a sample representative of the population (Burns & Burns, 2008). This experiment adopts a method of data collection that is based on the nature of the population. As such, the experiment is conducted in an online setting.

3.3.1 Sample

To some measure, many studies seek to analyse the entire defined population of their study. However, most of the time, due to the large size of these populations, that becomes extremely difficult or even impossible to do (Burns & Burns, 2008). In cases like these, researches have to rely on representative samples in order to make valid conclusions (Easterby-Smith, Thorpe & Jackson, 2015). The population of which this study is interested in are users of the online video search engine that is YouTube. The website currently has over one billion active users (YouTube, 2017b), making analysis of the entire population impossible in the given time frame. Therefore, due to the large size of the population, this study appoints a sample.

The sample of the population of interest is strategically narrowed down to YouTube users of 18-34 years of age. The sample is similar to the one used by Li and Lo (2015), who provide research suggesting that the said age range is YouTube’s dominant target group (ComScore, 2011). A more recent study by ComScore (2016) suggests the same findings, and therefore justifies our decision for choosing this age group.

3.3.2 Sampling Technique

The selection of the sampling technique influences the validity of the results. Due to its positive influence on validity, probability sampling is superior to non-probability sampling (Burns & Burns, 2008; Malhotra, 2010). The former, however, has disadvantages in that it requires substantial economic resources among other things (Bryman & Bell, 2011; Malhotra, 2010). Moreover, an essential condition for conducting probability sampling is being in possession of a sampling frame, a list of all units of the studied population (Burns & Burns, 2008; Malhotra, 2010). During the period of this study, it was not possible to receive a list of every YouTube user worldwide, and consequently, a sampling frame was not present. Therefore, a non-probability sampling technique was used for the purpose of this study.

The sampling technique chosen from the non-probability sampling family is convenience sampling. As the word implies, convenience sampling seeks experimental subjects from conveniently accessible points (Burns & Burns, 2008; Malhotra, 2010). The reasons for selecting the technique are that it is superior in terms of cost and time requirements (Malhotra, 2010). However, we are also aware of the limitations attached to this sampling technique that mostly circle around possible population misrepresentation. Since the selection of subjects is based on subjective selection, and not by numerical selection, selection bias is evident (Malhotra, 2010). Nonetheless, due to the immense size of the population and other constraints, convenience sampling was chosen as the most fitting sampling technique for this research.

3.3.3 Collection Method

Experiments often attract collection methods that are both timely and expensive, such as offline laboratory settings that require the presence of a large number of subjects at a given point in time. To counterbalance this limitation, our method of data collection utilises the advantages of the online environment, whereby subjects are limitless and unbounded by time. The experiment was conducted through a survey platform that allowed for the simulation of an experimental environment. The method is fitting for this type of quantitative research because it gathers data of a particular point in time that can later be used for statistical analysis with relative ease (Burns & Burns, 2008). The survey is of the structured type, meaning that questions are asked in a prearranged order and requires respondents to select answers from a predetermined set of responses (Malhotra, 2010). Due to this structure, the process is classified as direct, because to some extent it discloses the purposes of the research (Burns & Burns, 2008; Malhotra, 2010). As mild as the disclosure may be, it can potentially result in some degree of bias. However, the structured questionnaire compensates for it with other advantages. Due to its strict nature, it gathers reliable data, as responses are limited to the given answers. Also, the use of a survey simplifies the coding, analysis and interpretation of data in the later stages of the study (Malhotra, 2010).

3.3.4 Questionnaire

The survey is conducted electronically with its research tool being an online questionnaire. Apart from cost constraints, there are several reasons for conducting the survey in digital space. Firstly, the physical presence of interviewers is not required, because the questionnaire is self-administered (Burns & Burns, 2008), thus, solving time constraints. Moreover, online questionnaires are also more superior to their electronic survey alternative, an e-mail survey.

This is due to the fact that email surveys are often faced with compatibility issues or are perceived as spam (Burns & Burns, 2008). Also, unlike e-mail surveys, online questionnaires can contain graphics and sound (Malhotra, 2010), which is of crucial importance since the research studies and uses audio-visual content.

Of course, just like other data collection tools, online questionnaires come with their own limitations. Regardless of how people come across the online questionnaire, there is risk that unintentionally targeted subjects decrease the representativeness of the sample (Burns & Burns, 2008). Also, since the online questionnaire is self-administered, the response rate is hindered (Burns & Burns, 2008). Additionally, completion of the questionnaire was a definite limitation that was faced during the experiment, which may have been overcome if administered offline. Being aware of these limitations, we have taken measures to address them that mostly included adding numerous screening questions. The risk of low response rates was counterbalanced through certain measures suggested by Malhotra (2010), including reducing effort required, offering a legitimate purpose for each question, and avoiding sensitive information. Reducing effort was ensured by asking clear and short questions that were answered with either drop-down or Likert-scale options, while recall questions required no more than one sentence or less. Legitimacy was ensured by providing a short explanation where necessary, although most questions had a clear purpose. Questions that required a statement of purpose included the randomisation questions that were based on the month and day of birth. Here, a simple statement clarifying the need for randomisation was used. Additionally, sensitive information was completely avoided, other than birthday questions.

It is also important to mention that conducting the survey in digital space is fitting given the population of the study - those being online video users in the same digital space. Using an online survey allows the respondents to come in contact with the questionnaire while surfing the Internet in the comfort of their own homes. Considering that the top two reasons people watch YouTube videos are relaxation and entertainment (Google, 2016), this method is desirable in creating an accurate simulation.

3.3.5 Selection of Original YouTube Video

The original video was carefully selected using certain requirements such as (1) avoiding offensive content, (2) displaying high video quality, (3) having a high view count and like/dislike ratio to ensure the video is interesting to the target population and (4) having a length of approximately 2 minutes (Li & Lo, 2015). The last requirement was needed in order to avoid the experimental advertisement dominating too much of the original YouTube video, and at the same time not to consume too much of the respondents' time. On a separate note, video advertisements are shown in YouTube only if the original video is at least 30 seconds long (Google, 2017a). The chosen video was produced and posted online by BuzzFeed, an international digital media company that specialises in news and entertainment (BuzzFeed, 2017). With its target audience being people between 18 to 34 years old (BuzzFeed, 2017), the video channel provided the appropriate target population. More specifically, the video shown in each of the four experiment groups was titled "People Try the Spicy Taco Challenge". The video was of high quality (720p resolution), had over 1 million views and a highly acceptable like/dislike ratio of 34 to 1. Apart from satisfying the initial requirements, the video was also chosen because of its universally relatable theme, food, as well as its entertaining nature that YouTube users are fond of (Google, 2016).

3.3.6 Selection of Video Advertisement

Criteria for choosing a video advertisement for the experiment was as follows: (1) to have both 5- and 15-second versions that contain the same content, theme, message and only differ in length, (2) to be displayed in high video quality, (3) to have limited vocals, (4) to have clear exposure of the brand logos, (5) to advertise a brand that is relatively unknown, and (6) not to contain celebrity endorsement (Li & Lo, 2015). The last two requirements were necessary in order to avoid familiarity with the promoted brand and those promoting it, since it can influence how viewers process the advertisement (Dens & De Pelsmacker, 2010). The chosen video advertisement promotes a Mexican tequila brand named ‘Don Weber’, a brand that is unknown in the European market. The advertisement had two similar versions of which were exactly 5- and 15-seconds long, both being of high quality (720p resolution), and containing the same background music. Both of the advertisements also had identical verbal and written messages that mentioned the brand name ‘Don Weber’ and its slogan ‘simply the best’.

Once choosing both the original YouTube video and the advertisement, another important factor considered was the congruity between the two videos. This was an important consideration given that YouTube places advertisements in YouTube videos according to certain keywords or themes, of which are congruent with the original videos (Google, 2017b). Since the chosen YouTube video provides content of spicy tacos, a traditional Mexican dish, its combination with an advertisement of tequila, a traditional Mexican beverage, is congruent, and thus, satisfies the last requirement of the video.

In order to simulate a natural YouTube environment, the combined videos were edited by adding additional visual effects to the advertising part of the video. These effects involved carefully replicating banner type animations (Figure 3b) that appear when advertisements are played during a video on YouTube (Figure 3a). This increased the authenticity of the simulation and ensured subjects did not mistake the advertisement as a part of the original video.



(a) Original YouTube Video



(b) Advertisement with banner animations

Figure 3. Screenshots of video advertisement and YouTube video used in the experiment

3.3.7 Pretest

Pretesting was used on a small sample to test for potential problems of the experimental design. Respondents of the pretest were drawn from the same population as the main study and included a total of twelve participants. Although the main study is conducted online in an anonymous fashion, the pretest was conducted face-to-face, although still using the online

survey template. This was done so that attitudes regarding the survey could be directly observed. Of particular interest to test was the wording, question difficulty, instructions and question interpretation. The pretest resulted in adapting the wording to more simplistic vocabulary for the international sample. This also resolved issues regarding question difficulty. The question measuring product familiarity was found to be interpreted differently than expected, which was altered by adding an explanation. During the pretest, protocol analysis was adopted (Malhotra, 2010), which involved asking respondents to “think aloud” when completing the survey. After the survey had been completed, respondents were asked to describe what each question meant, as well as if they encountered any issues during the survey. The same twelve respondents were asked to complete the survey one week later, in order to measure the reliability of the measurement scales that are utilising Likert scales. Results were placed into the statistical analysis software SPSS, and the correlation coefficient was computed to test whether deviation in results occurred. This part of the procedure is further explained in the Validity and Reliability section of the Methodology.

3.3.8 Procedure

The questionnaire was created and launched on the online survey platform, Survey Monkey. The said platform was chosen due to its compatibility with audio-visual content and additional features that helped to randomly assign respondents to different experiment groups. Due to the fact that the sample contained people of various nationalities, the chosen language was English. The questionnaire was distributed using Facebook, the world’s largest social media network (Statista, 2017b), through which we had the biggest reach in comparison to other platforms. The survey was launched and actively gathered data for a total of seven days (April 15th - April 21st, 2017). The reason for choosing such a timeframe was to cover all seven days of the week to reduce the ‘day of the week effect’, given that some people only surf during the weekend or more so during particular days of the week (Burns & Burns, 2008).

Once opening the survey, respondents were shortly introduced to the objective of the study without being exposed to the specifics, in order to avoid bias. After answering demographics related questions, all underage respondents were screened out due to the fact that YouTube does not allow for alcohol advertisements targeting underage audiences (Google, 2017c). After identifying the information needed for the previously mentioned randomisation, the respondents were assigned to four experiment groups, differing only by the length of the advertisement and its placement within the video. Questions both prior and after the video remained the constant across experimental groups. The completion of the survey average 3-4 minutes.

Experimental studies such as this one require all fragmented experimental groups to contain similar subjects (Burns & Burns, 2008). For example, there should be a similar level of intelligence among each group. In order to encourage this, we have adopted a recommendation made by Burns and Burns (2008), whereby subjects are randomly assigned to experimental groups using a combination of their month and day of birth. If both values are even, subjects are assigned to experiment group 1; if subjects were born on an even month and an odd day, they are assigned to experiment group 2; if subjects were born in an odd month and an even day, they are assigned to experiment group 3 and so forth. In total, there were four experiment groups that contain a variation of advertisement length and interruption. See Table 1 for a visual presentation of the assignment of advertisement to experimental groups.

Group	Advertisement Length	Placement
Group 1	5-seconds	Pre-roll
Group 2	5-seconds	Mid-roll
Group 3	15-seconds	Pre-roll
Group 4	15-seconds	Mid-roll

Table 1. Experimental groups and assigned advertisements

After watching the videos, the respondents faced a second screening question, asking them about the content of the video. The purpose of this question was to screen out all the people that did not finish the entire video, which is a necessity given that some advertisements were placed midway into the video, and if skipped, would decrease the credibility of the following answers. Afterward, the respondents were asked whether they were users of YouTube, more specifically, whether they have seen a YouTube video in the past year. The purpose of this question was to screen out all subjects who were not a part of the tested population. Additionally, subjects who have not been using YouTube recently may not be familiar with the banner type animations that YouTube displays when showing advertisements. A third screening question involved asking participants whether they noticed an advertisement at all. If the answer was no, the respondents were screened out because of their possible incapability to answer advertisement-related questions that followed. Subjects who had seen the advertisement before the experiment were also screened out, given their significantly better chances to recall the brand name and message. That would significantly decrease the homogeneity of the respondents, and thus, the credibility of the findings.

After successfully going through the screening questions the respondents were asked questions that would measure brand recall, message recall, the likeability of the advertisement, as well as the familiarity with the advertised product category. These questions and the measurements of their answers are discussed in more detail in the following section.

Before the subjects qualified for statistical analysis, they were asked a final screening question, which questioned whether they were familiar with the brand ‘Don Weber’ prior to seeing the video advertisement. The reason for asking this particular question in the very end of the questionnaire was not to jeopardise the recall questions, since the screening question could have given the answer away. See Table 2 on the following page for a graphical representation of the ordering of questions and their purpose.

No.	Question	Purpose
1.	Where are you from?	Demographics
2.	What is your gender?	Demographics
3.	How old are you?	Demographics and screening
4.	In which month were you born?	Randomisation
5.	On which day were you born?	Randomisation
The video		
6.	Did the people in the picture finish eating their tacos?	Screening
7.	Have you seen this video before?	Screening
8.	Have you watched a video on YouTube in the past year?	Screening
9.	While watching the video did you encounter an advertisement?	Screening
10.	Have you seen this advertisement before?	Screening
11.	What is your opinion of the advertisement?	Measurement of likeability
12.	What is the name of the tequila brand shown in the advertisement?	Measurement of brand recall
13.	What is the brand's slogan used in the tequila advertisement?	Measurement of message recall
14.	How familiar are you with tequila as a product?	Measurement of familiarity
15.	Have you heard of the brand 'Don Weber' before?	Screening

Table 2. Survey questions and their purpose

3.3.9 Measurements

Measurement and scaling are two of the most important procedures that researchers decide on and must be as accurate as possible (Malhotra, 2010). Accuracy was increased in this study by adopting measurements from previous research; the 5-point scale was used to measure advertisement likeability (Haley & Baldinger, 2000; Jeong, Sanders & Zhao, 2011; Newstead & Romaniuk, 2009); open ended questions with an aided product category prompt were used to measure both brand and message recall (Danaher & Mullarkey, 2003; Haley & Baldinger, 2000; Newstead & Romaniuk, 2009; Singh & Cole, 1993); and a 5-point scale was used to measure product familiarity (Bettman & Park, 1980; Johnson & Russo, 1984).

Advertisement likeability was measured by asking participants a single question with a 5-point scale. The scale was adopted from previous research by Haley and Baldinger (2000), Jeong, Sanders and Zhao (2011), and Newstead and Romaniuk (2009) who also tested for likeability regarding advertisements. This justified our choice for a single item scale. The question involved participants choosing the most appropriate statement that best described their attitude towards the advertisement (*1= I disliked it very much, 2= I disliked it, 3= I neither liked nor disliked it, 4= I liked it, 5= I liked it very much*). This type of interval-level scale assigns numbers to objects so that they can be compared (Malhotra, 2010). An advantage of such a measurement is that the arithmetic mean and standard deviation can be calculated (Malhotra, 2010). Itemised rating scales, such as this one, are widely used in marketing research (Malhotra, 2010). For analysis, the data can be compared on an item-by-item basis, or a summed score can be calculated (Malhotra, 2010). This sort of measurement was suitable for the study because (1) it is easily understood by respondents, ensuring more accurate data, and (2) its simplicity is suitable for online surveys with no explanation required. However, a disadvantage of scaling is that respondents require more time to respond to the question because they must read and evaluate each statement individually (Malhotra, 2010). This disadvantage may be amplified in an online experimental environment where respondents do not wish to invest too much time in their answers.

Brand and Message recall were measured by asking participants an open question that provided aided material in disclosing the advertised product category, adopted from Danaher and Mullarkey (2003), Haley and Baldinger (2000), Newstead and Romaniuk (2009), as well as Singh and Cole (1993). The results were evaluated using a dichotomous variable (correct choice = 1; incorrect choice = 0). This type of nominal scale serves to classify results based on a one-to-one resemblance between the number and the object and is mutually exclusive (Malhotra, 2010). Nominal measurements such as these are limited to inferential statistics that use frequency counts.

Control variables used for this study, as well as other studies, were product familiarity (Bettman & Park, 1980; Johnson & Russo, 1984), age (Danaher & Mullarkey, 2003; Li & Lo, 2015), gender (Danaher & Mullarkey, 2003; Li & Lo, 2015) and nationality (Li & Lo, 2015). Product familiarity was measured by asking participants a single question with a 5-point scale, adopted from Bettman and Park's (1980), and Johnson and Russo's (1984) studies on advertisement effectiveness. The question involved participants choosing the most appropriate statement that best described their familiarity with the product category (*1= very unfamiliar, 2= unfamiliar, 3= neither familiar nor unfamiliar, 4= familiar, 5= very familiar*). Age, gender and nationality were all measured on the nominal basis.

An overview of the measured variables is depicted in the table below (Table 3). Each variable is categorised by independent, dependent and control variables. In addition, the variables are labelled in accord to the scale they are measured by.

Measureable Variables	Type of Variable	Measurement
Advertisement Length	Independent variable	Nominal
Brand Recall	Dependent variable	Nominal
Message Recall	Dependent variable	Nominal
Advertisement Likeability	Dependent variable	Interval
Interruption	Moderating variable	Nominal
Product Familiarity	Control variable	Interval
Age	Control variable	Nominal
Gender	Control variable	Nominal
Nationality	Control variable	Nominal

Table 3. Categorisation of variables used in the experiment

3.4 Data Processing

After data collection was finalised, the study had collected 271 responses via the online questionnaire. This data was prepared in terms of editing, coding and transcribing before analysis began (Malhotra, 2010). The initial step in this process concerned questionnaire checking. During this step, the screening questions that were mentioned in the data collection section were used to eliminate 57 responses of non-targeted participants, for example those below the age of 18 and above the age of 34. The second step in the data processing involves editing, whereby 20 incomplete questionnaires were discarded (Malhotra, 2010). After this step, the study was left with a total of 194 representative responses. This resulted in each of the experimental groups containing over 30 respondents, a size and distribution considered

reliable for statistical analysis (Wilson Van Voorhis & Morgan, 2007). The respective counts for each of the four experiment groups were 44, 53, 45 and 52. Of these 194 total participants, most were Lithuanian (75,7%), followed by German (8,7%). The remaining 15,6% of participant nationalities consisted of American, Australian, Brazilian, Czech, Finnish, Italian and Swedish. Regarding age, 72,1% of respondents were aged 18-24, 21,6% aged 25-29 and 6,3% aged 30-34. Regarding gender, there was a rather even split between males (45,8%) and females (54,2%). Refer to Table 4 for an overview of the sample demographics.

Control Variables		Count	%
Age	18-24	140	72,1%
	25-29	42	21,6%
	30-34	12	6,3%
	Total	194	100%
Gender	Male	89	45,8%
	Female	105	54,2%
	Total	194	100%
Nationality	Lithuanian	146	75,5%
	German	17	8,7%
	Other	31	15,6%
	Total	194	100%

Table 4. Overview of the sample demographics

Data was then coded, in which numbers were assigned to responses depending on their answer (Malhotra, 2010). Regarding the measurements involving Likert scales, answers were pre-coded, from 1 to 5. Regarding recall measurements, answers were translated into a dichotomous measure (correct recall = 1; incorrect recall = 0). Challenges occurred when coding recall data, as spelling errors and incomplete answers distort reliability. Small deviations from the correct answer were accepted, given that the sample were not native English speakers. Independent variables received numbers corresponding to length (5-second = 0; 15-second = 1) and interruption (uninterrupted, pre-roll = 0; interrupted, mid-roll = 1). In terms of the control variables, familiarity was coded similarly to likeability, on a 5-point scale, with 1 = very unfamiliar and 5 = very familiar. Age was coded into four main groups; 1 = 18-24, 2 = 25-29, 3 = 30-34. Gender was coded as 0 = male and 1 = female. After observing the data, Nationality was split between Lithuanian and non-Lithuanian given the large percentage of the former. Data was then conveniently transcribed into SPSS from Survey Monkey's platform given the compatibility between the two platforms.

Once editing, coding and transcribing was finalised, statistical analysis was performed using SPSS software (version 24). The strategy for analysis was selected on the basis of the data derived from the causal research design. The characteristics of the data strongly influence the statistical tests that are used, whether univariate or multivariate (Malhotra, 2010). The study aims to examine the relationship between the advertisement length and each dependent variable in isolation, satisfying the criteria for univariate techniques (Malhotra, 2010). Univariate techniques can be further differentiated according to the measurement scale, whether metric or nonmetric (Malhotra, 2010). Metric data is measured using interval or ratio scales, and nonmetric data is measured using nominal or ordinal scales (Malhotra, 2010). Our dependence variables are split in this respect, whereby the measurement of advertisement likeability is classified as metric, while the measurements of both brand and message recall are nonmetric.

Based on the above criteria, the fitting method of analysis for brand and message recall was loglinear analysis, while the method of analysis for advertising likeability was initially analysis of covariance (ANCOVA). Both statistical tests were selected based on the nature of independent, dependent and control variables they analyse (Field, 2009). However, after observing the distribution of the control variables through ocular inspection, it was deemed unnecessary to include them in statistical analysis. Table 5 illustrates the said equal distribution among the four experiment groups. Age between the experimental groups was approximately 70% from 18-24, 20% from 25-29, and 6% from 30-34. In terms of gender, there was a rather equal distribution between males and females, with close to 50% of each gender in each group. Nationality was the most unevenly distributed among the control variables, but did not raise too much concern, since the values were still relatively even, with approximately 75% Lithuanian and 25% non-Lithuanian in each group. Familiarity with the product averaged a value of approximately 2,50 between each experimental group. Additionally, a correlation matrix was performed, as shown in Appendix A. No significant correlations were revealed between each of the dependent variables and the control variables. As one assumption of ANCOVA requires correlations to occur between the dependent and control variables, this assumption was violated (Field, 2009). Consequently, the adjusted statistical test for advertisement likeability was two-way analysis of variance (ANOVA); brand and message recall remained with loglinear analysis given that it did not depend on the use of control variables (Burns & Burns, 2008; Field, 2009).

Control Variables		Group 1	Group 2	Group 3	Group 4
Age	18-24	75%	74%	71%	71%
	25-29	18%	21%	22%	25%
	30-34	7%	6%	7%	4%
	Total	100%	100%	100%	100%
Gender	Male	48%	43%	47%	48%
	Female	52%	57%	53%	52%
	Total	100%	100%	100%	100%
Nationality	Lithuanian	80%	75%	78%	73%
	Non Lithuanian	20%	25%	22%	27%
	Total	100%	100%	100%	100%
Familiarity	Means (scale: 1-5)	2,66	2,42	2,43	2,60

Table 5. Overview of distribution of control variables

3.4.1 Assumptions of Loglinear Analysis

Loglinear analysis is a form of regression used to test significant relationships between three or more nominal variables (Field, 2009). It provided an analysis of the effect of advertisement length and interruption on nominal dependent variables, those being brand and message recall. Separate loglinear analyses were used for each of the recall variables, because of the univariate nature of the analysis (testing variables in isolation). The three nominal variables include advertisement length, interruption and each of the recall variables separately. It might also be important to recognise that there is no distinction between dependent and independent variables in loglinear analysis, as they are treated equally (Field, 2009). However, given the setup of the research study, the outcome variables are known to the authors, those being brand or message recall. Loglinear's assumptions require that observations are independent and expected frequencies are large enough to produce reliability. The latter assumption is met

when (1) 80% or more of the cells have expected frequencies equal or greater than 5 and (2) all cells have an expected frequency greater than 1 (Field, 2009). If one or both of the assumptions are violated, then there will be a significant loss in statistical power (Field, 2009). After testing for the said assumptions, all of them were met, as seen in Table 6 and Table 7. Brand recall retained all frequencies above 5, while message recall resulted in 12,5% of cells below the expected frequency of 5. However, this was still below the critical level of 20%.

Ad Length	Interruption	Brand recall	Observed		Expected		Residuals	Std Res.
			Count	%	Count	%		
5-second	Not Interrupted	Incorrect	38,500	19,8%	38,500	19,8%	0,000	0,000
		Correct	6,500	3,4%	6,500	3,4%	0,000	0,000
	Interrupted	Incorrect	45,500	23,5%	45,500	23,5%	0,000	0,000
		Correct	8,500	4,4%	8,500	4,4%	0,000	0,000
15-second	Not Interrupted	Incorrect	39,500	20,4%	39,500	20,4%	0,000	0,000
		Correct	6,500	3,4%	6,500	3,4%	0,000	0,000
	Interrupted	Incorrect	28,500	14,7%	28,500	14,7%	0,000	0,000
		Correct	24,500	12,6%	24,500	12,6%	0,000	0,000

Table 6. Brand recall: Testing the assumptions of loglinear analysis

Ad Length	Interruption	Message recall	Observed		Expected		Residuals	Std Res.
			Count	%	Count	%		
5-second	Not Interrupted	Incorrect	40,500	20,9%	40,500	20,9%	0,000	0,000
		Correct	4,500	2,3%	4,500	2,3%	0,000	0,000
	Interrupted	Incorrect	47,500	24,5%	47,500	24,5%	0,000	0,000
		Correct	6,500	3,4%	6,500	3,4%	0,000	0,000
15-second	Not Interrupted	Incorrect	38,500	19,8%	38,500	19,8%	0,000	0,000
		Correct	7,500	3,9%	7,500	3,9%	0,000	0,000
	Interrupted	Incorrect	39,500	20,4%	39,500	20,4%	0,000	0,000
		Correct	13,500	7,0%	13,500	7,0%	0,000	0,000

Table 7. Message recall: Testing the assumption of loglinear analysis

3.4.2 Assumptions of Two-Way ANOVA

The second statistical test used in SPSS is two-way ANOVA. The test provides an analysis of the effect of advertisement length and interruption on dependent variables measured on the interval scale, that being advertisement likeability.

Before conducting the two-way ANOVA, multiple assumptions were met. For one, the dependent variable was measured at the continuous level, in which case likeability was measured using an interval scale. Secondly, both the independent and moderating variables had two or more categorical, independent groups. Advertisement length was categorised as 5- and 15-second, while interruption was categorized as pre-roll and mid-roll. Thirdly, the observations in each group were independent of each other, meaning each group contained mutually exclusive subjects.

Three additional assumptions were tested using SPSS, including testing for significant outliers, normal distribution and homogeneity of variance (Burns & Burns, 2008). No outliers

were present in the data after observing the box plots produced by SPSS (see Appendix B). Additionally, the associated probability of the Levene's test statistic shows homogeneity of variance ($p = 0,695$), accepting the null hypothesis that error variances of the dependent variable are equal across groups. In terms of testing for normal distribution, the Shapiro-Wilk test showed significance ($p < 0,001$), leading to the rejection of the null hypothesis that the data is normally distributed. This rejection may be due to the fact that the test is sensitive to sample sizes. Thus, we rely on a visual inspection before making any conclusion about normality. For validation, the histograms in Figure 4 were observed for normal distribution. The resulted outcome was a fairly normal distribution. Given that ANOVA is fairly robust to violations of normality, and the study hosts a large sample size ($N = 194$), the analysis can proceed with accurate results regardless of the violation of normality. The robustness of normality is based on the central limit theory, which suggests that the distribution of the mean of data will approach normality with the presence of a large sample size.

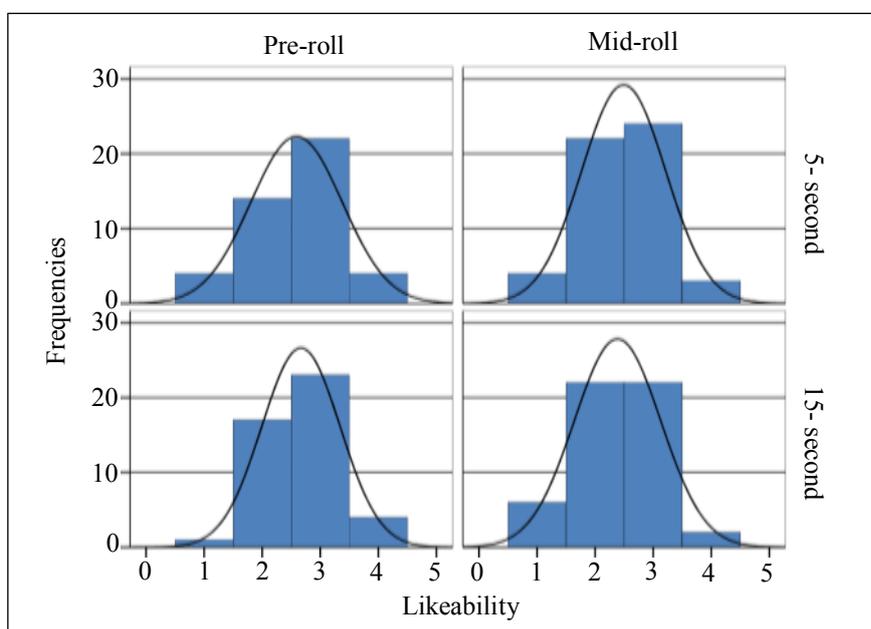


Figure 4. Advertisement likeability: Testing the assumptions of two-way ANOVA

3.5 Validity and Reliability

Validity in experimentation is of extreme importance in order to draw valid conclusions and make generalisations about the population (Malhotra, 2010). Malhotra (2010) refers to two types of validity - internal and external. The former concerns whether the findings reflect the true relationship between the independent and dependent variables. Extraneous variables must be controlled for in order to make valid conclusions about the relationship between the tested variables. The latter of the two concerns the generalisability of the findings. Malhotra (2010) suggests that it may not be valid to draw generalisations of a population when the experiment has not considered other extraneous variables that usually alter the relationship between independent and dependent variables.

Aspects of validity were mentioned in the previous sections, but will be discussed in further detail here. Often in marketing research, researchers must strike a balance between internal and external validity (Malhotra, 2010). The experiment was conducted in a controlled semi-laboratory setting, as opposed to a field setting. Reasons for this was to increase control, internal validity and ease of implementation as well to combat the limitations of time and budget (Malhotra, 2010). This decision, however, produced some limitations. For one, the controlled environment is artificial which results in lower external validity and is high in reactive error (Malhotra, 2010). Reactive error refers to the situation where subjects respond to stimuli in an artificial way, because of the knowledge that they are part of an experiment (Malhotra, 2010). This is similar to demand artefacts, which concern respondents answering questions based on what they believe to be the wanted answer (Malhotra, 2010). This was avoided in two ways; recall questions avoided demand artefacts all together because the nature of the question is built from memory, rather than subjective response; and subjects were diverted from the actual experiment by implying questions will be based on the YouTube video as well. This latter method increased validity in itself because it stimulated a more realistic environment whereby viewers of online videos pay more attention to the original video rather than the advertisement. Furthermore, a laboratory-like settings, as opposed to a field setting, is often superior as it allows for more complex experimental designs. This in turn allows for the control of many variables that may increase the validity of the results.

An important factor that may have distorted reliability of the results is the fact that mostly non-native English-speaking subjects were recruited for the experiment, although the experiment was in English. This may decrease validity because of language limitations that play a role in the processing and recall of information. This limitation was difficult to overcome given that the authors had a scarce network of native English speaking subjects to recruit for the experiment.

The validity of the measurement and scaling used within the experiment depend on many factors, one of which includes measurement accuracy (Burns & Burns, 2008). Measurement accuracy involves the observed score equalling the true score (Burns & Burns, 2008; Malhotra, 2010). In reality, the observed score includes a fraction of the true score in addition to systematic and random error (Malhotra, 2010). Systematic error concerns factors that affect the observed score in a constant way, while random error concerns situational factors that differ each time a measurement is made (Malhotra, 2010). Reliability and validity are mostly concerned with random error. Scale measurements are reliable when repeated measurements produce consistent results, in other words, when results are free from random error (Malhotra, 2010). A method used to test the reliability of a scale, according to Malhotra (2010), is the test-retest method. This study adopted Malhotra's (2010) recommendation by testing twelve respondents twice, in similar conditions, in order to calculate the degree of similarity between the two responses. The correlation coefficient calculated during the pre-test was 0,607 ($p = 0,036$) for advertisement likeability, assuring reliability. However, due to time restrictions, the recommended time of two weeks between tests (Malhotra, 2010) could not be obliged. Instead, one week was deployed between tests. This was a limitation that could not be avoided but still provided valuable insights into the reliability of the results. A separate limitation of this method is that responses may genuinely change over a span of one week, irrespective of random error. However, it is important to note that the above test was complementary to other tests that have proven the reliability of these measurements (Haley & Baldinger, 2000).

3.6 Ethical Considerations

There continues to be an crucial dilemma in experimental research design - the balance between ethical principles and reliable or valid research (Burns & Burns, 2008). The purpose of any experiment should be disguised in order to reduce demand artefacts, however, this is often considered unethical given the deception involved (Malhotra, 2010). Making informed consent absolutely essential, which includes providing information about the nature of the experiment, was introduced after the Nuremberg war trials (Burns & Burns, 2008). Since disguising the true purpose of the experiment increases response bias, it should be avoided in such a way that does not violate the rights of the respondents (Malhotra, 2010). This was done in several ways within the experiment; subjects were informed that the experiment was disguised before they participated in the experiment; they were also told that they could quit the experiment at any time; and the true purpose of the experiment was revealed at the end of the survey, for which they were able to comment if they did not want their results to be used. This is also referred to as debriefing and ensuring a reduction in response bias when conducted after the experiment (Malhotra, 2010).

The right to anonymity and confidentiality is another important ethical consideration in research (Burns & Burns, 2008). This was ensured during the experiment in two ways; participants' answers were mostly anonymous; however, IP addresses were used to eliminate multiple entries from the same source in order to increase reliability; and all information provided was strictly confidential, accessed only by the authors of the paper. Anonymity and confidentiality were ensured in the introduction text of the survey before the experiment was conducted. Burns and Burns (2008) also mention an important phenomenon that has affected ethics in research, which is the impact of technology. This has been a relevant issue in our research as the experiment was conducted using an online survey. One issue that technology induces, as proposed by Burns and Burns (2008), is that the relationship between participants and the researcher is 'faceless'. This produces its own issues since the researchers are often unaware if the sample is suitable (i.e. if children participate). This was another important issue that needed to be addressed in this experiment, since the advertisement concerned alcohol, and must not be promoted to children. To overcome this, the survey included a filtering question, which disqualified participants who were under the age of 18 based on the average worldwide drinking age. However, it is important to note that the risk of minors participating in the survey is never completely eliminated with anonymous online surveys.

4 Results

4.1 Results for Brand Recall

As previous research has shown, there is a strong positive relationship between advertisement length and the recall of the brand it promotes (Danaher & Mullarkey, 2003; Haley & Baldinger, 2000; Newstead & Romaniuk, 2009). Previous research has also shown that interruption strengthens this positive relationship (McCoy et al., 2008). Respectively, the first two hypotheses of this study are testing whether such relationships are present in the online environment by comparing the 5-second and 15-second video advertisements.

***H1a:** 15-second online video advertisements generate better brand recall than 5-second online video advertisements.*

***H2a:** Interruption strengthens the positive relationship between online video advertisement length and brand recall.*

A loglinear analysis was conducted in order to test the relationship between advertisement length and brand recall and the effect that interruption has on it. The results of loglinear analysis provide three main effects (length, interruption, recall), as well as three interactions effects of two variables (length x interruption, length x recall, interruption x recall) and one interaction effect of three variables (length x interruption x recall). The analysis begins by including all effects and then removes the highest-order interaction after which a new model is constructed (Field, 2009). This is known as backward elimination. From the model, the expected frequencies are calculated and are compared with the observed frequencies using the likelihood ratio statistic (Field, 2009). The effect of the interaction can be measured by observing whether there is a significant change in this statistic with the removal of the interaction. If so, a significant interaction exists between the given variables. If not, a new model is constructed by removing lower-order interactions (Field, 2009). The cell counts and residuals table (Table 8) presents the observed and expected counts for each of the three-way interactions (length x interruption x recall).

Ad Length	Interruption	Brand recall	Observed		Expected		Residuals	Std Res.
			Count	%	Count	%		
5-second	Not Interrupted	Incorrect	38,500	19,8%	38,500	19,8%	0,000	0,000
		Correct	6,500	3,4%	6,500	3,4%	0,000	0,000
	Interrupted	Incorrect	45,500	23,5%	45,500	23,5%	0,000	0,000
		Correct	8,500	4,4%	8,500	4,4%	0,000	0,000
15-second	Not Interrupted	Incorrect	39,500	20,4%	39,500	20,4%	0,000	0,000
		Correct	6,500	3,4%	6,500	3,4%	0,000	0,000
	Interrupted	Incorrect	28,500	14,7%	28,500	14,7%	0,000	0,000
		Correct	24,500	12,6%	24,500	12,6%	0,000	0,000

Table 8. Brand Recall: Cell counts and residuals

The K-way and higher order effects table (Table 9) is divided into two parts. The first row (K = 1) in the ‘K-way and higher order effects’ section tells us that if removing the one-way effects (the main effects of length, interruption and brand recall), and any of the higher order effects (length x interruption, length x recall, interruption x recall, length x interruption x recall), will significantly affect the model fit. In other words, K = 1 tests if the fit of the model will change if all the factors were removed (Field, 2009). The removal of these result in statistically significance ($p < 0,001$). The second row (K = 2) removes the two-way effects (length x interruption, length x recall, interruption x recall) as well as the three-way effect (length x interruption x recall). This is also statistically significant ($p < 0,001$). K = 3 also proves significance ($p = 0,041$) when removing only the three-way effect.

The ‘K-way effects’ section looks at the one-way effect (K = 1), two-way effect (K = 2) and three-way effect (K = 3), without considering the higher order effects like the first section did. Therefore, K = 1 determines if the model would be significantly affected if only the main effects were removed. This yields a significant result ($p < 0,001$). Removing only the two-way effects, as tested in K = 2, the results also show statistical significance ($p = 0,001$). K = 3 produces a significant result ($p = 0,041$) after removing the three-way effect, which shows that the three-way interaction is also a strong predictor.

The tables thus far have presented results indicating a significant interaction ($p = 0,041$) between advertising length, interruption and brand recall, which is relevant to the hypothesis concerning interruption as a moderating variable. Additionally, the two-way effects yield significant results ($p < 0,001$). However, one must look at the partial associations in order to specify exactly which interaction it concerns.

	K	df	Likelihood Ratio		Pearson		Iterations
			Chi-Square	Sig.	Chi-Square	Sig.	
K-way and Higher Order Effects	1	7	83,132	<0,001	73,464	<0,001	0
	2	4	20,599	<0,001	22,256	<0,001	2
	3	1	4,188	0,041	4,409	0,036	3
K-way Effects	1	3	62,533	<0,001	51,208	<0,001	0
	2	3	16,411	0,001	17,847	<0,001	0
	3	1	4,188	0,041	4,409	0,036	0

Table 9. Brand Recall: K-Way and higher-order effects

If we look at the partial associations (Table 10), length x interruption is not statistically significant ($p = 0,467$), but length x brand recall and interruption x brand recall are both statistically significant ($p = 0,004$; $p = 0,003$). Therefore, two of the three two-way interactions have statistical significance. Of interest to the hypothesis is the interaction between advertisement length and brand recall.

Effect	df	Partial Chi-Square	Sig.	Iterations
Length*Interruption	1	0,529	0,467	2
Length*Brand recall	1	8,171	0,004	2
Interruption*Brand recall	1	8,726	0,003	2
Length	1	0,000	1,000	2
Interruption	1	1,321	0,250	2
Brand recall	1	61,211	<0,001	2

Table 10. Brand recall: Partial associations

Taking a look at parameter estimates in Table 11, an important part of the table is the z-score, whereby all the different effects can be compared on a standardized score. As can be seen, the main effect of brand recall has the largest absolute value ($z = 7,140$). Also fairly strong would be interruption x brand recall ($z = 2,315$) and length x interruption x brand recall ($z = -2,048$), as well as length x brand recall ($z = 1,980$).

Effect	Estimate	Std. Error	Z	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
Length*Interruption*Brand Recall	-0,194	0,095	-2,048	0,041	-0,380	-0,008
Length*Interruption	0,071	0,095	0,746	0,456	-0,115	0,256
Length*Brand Recall	0,188	0,095	1,980	0,048	0,002	0,373
Interruption*Brand Recall	0,219	0,095	2,315	0,021	0,034	0,405
Length	-0,077	0,095	-0,813	0,416	-0,263	0,109
Interruption	-0,179	0,095	-1,894	0,058	-0,365	0,006
Brand Recall	0,677	0,095	7,140	<0,001	0,491	0,862

Table 11. Brand recall: Parameter estimates

After elimination, the final output produces the likelihood ratio statistic, based on the new model shown in Table 12. It tests whether the expected frequencies generated by the model are significantly different from the actual frequencies. In that sense, it is favourable for the tests to be non-significant, as they are in this case (Field, 2009).

Ad Length	Interruption	Brand Recall	Observed		Expected		Residuals	Std Res.
			Count	%	Count	%		
5-second	Not Interrupted	Incorrect	38,000	19,6%	38,000	19,6%	0,000	0,000
		Correct	6,000	3,1%	6,000	3,1%	0,000	0,000
	Interrupted	Incorrect	45,000	23,2%	45,000	23,2%	0,000	0,000
		Correct	8,000	4,1%	8,000	4,1%	0,000	0,000
15-second	Not Interrupted	Incorrect	39,000	20,1%	39,000	20,1%	0,000	0,000
		Correct	6,000	3,1%	6,000	3,1%	0,000	0,000
	Interrupted	Incorrect	28,000	14,4%	28,000	14,4%	0,000	0,000
		Correct	24,000	12,4%	24,000	12,4%	0,000	0,000

Table 12. Brand recall: Cell counts and residuals based on the adjusted model

4.1.1 Interpretation of Brand Recall Results

When interpreting the results, it is useful to plot the different frequencies for each of the categories. This is done by conducting a cross-tabulation (Field, 2009), as seen in Table 13 on the following page, which presents the frequencies of correct vs incorrect brand recall in each experiment group (5-second pre-roll, 5-second mid-roll, 15-second pre-roll and 15-second mid-roll). In addition, a graphical presentation of the frequencies in respect to brand recall is shown in Figure 5 on the following page.

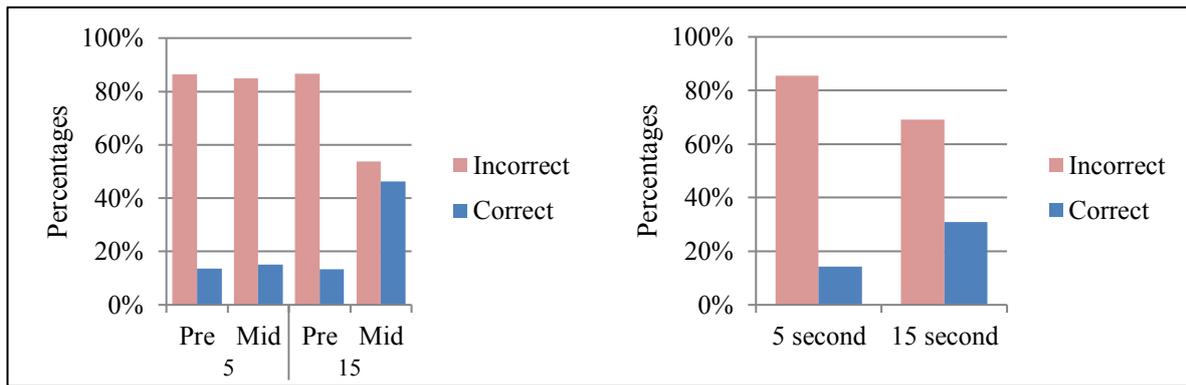


Figure 5. Brand recall: Graphical presentation of frequencies

In terms of hypothesis H1a, which tests the relationship between advertisement length and brand recall, the right graph in Figure 5 displays the superiority of 15-second advertisements. 15-second advertisements are more than twice as effective as 5-second advertisements in terms of brand recall, with 30,9% vs 14,4% recall respectively.

In terms of hypothesis H2a, which tests whether interruption strengthens the positive relationship between advertisement length and brand recall, the left graph in Figure 5 shows the superiority of interruption-based advertisements. This superiority is visible for both 5- and 15-second advertisements; however it is most recognised for 15-second advertisements. The results of Table 13 indicate that brand recall increased a mere 1,5% for 5-second advertisements while brand recall for 15-second advertisements significantly increased by 32,9%.

Interestingly, 15-second versions are superior to their shorter counterparts only when interrupted, while uninterrupted 15-second advertisements are inferior to both versions of 5-second advertisements.

Ad Length	Interruption	Brand Recall	Count	% within group
5-second	Not Interrupted	Incorrect	38	86,4%
		Correct	6	13,6%
	Interrupted	Incorrect	45	84,9%
		Correct	8	15,1%
	Total	Incorrect	83	85,6%
		Correct	14	14,4%
15- second	Not Interrupted	Incorrect	39	86,7%
		Correct	6	13,3%
	Interrupted	Incorrect	28	53,8%
		Correct	24	46,2%
	Total	Incorrect	67	69,1%
		Correct	30	30,9%

Table 13. Brand recall: Cross-tabulation with frequencies

In summary, a loglinear analysis has revealed statistically significant relationships between all tested variables in respect to hypotheses H1a and H2a. As a result of this conclusion:

→ We accept the hypothesis H1a and conclude that the length of online video advertising has a positive effect on brand recall.

→ We accept the hypothesis H2a and conclude that the interruption strengthens the positive relationship between online video advertising and brand recall.

4.2 Results for Message Recall

According to previous literature, a strong positive relationship exists between advertisement length and recall of its message (Fabian, 1986; Mord & Gilson, 1985; Singh & Cole, 1993). As with the before tested hypotheses, previous research has shown that interruption strengthens this relationship as well (Moorman, Neijens & Smit, 2002). Analogical to the previously tested hypotheses, the next two hypotheses test whether such relationships are present in the online environment by comparing the 5-second and 15-second video advertisements.

H1b: 15-second online video advertisements generate better message recall than 5-second online video advertisements.

H2b: Interruption strengthens the positive relationship between online video advertisement length and message recall.

A loglinear analysis was conducted in order to test the relationship between advertisement length and message recall as well as the effect that interruption has on this relationship.

As with the previous section, the cell counts and residuals table presents the observed and expected counts for each of the three-way interactions (length x interruption x recall), as shown in Table 14.

Ad Length	Interruption	Message recall	Observed		Expected		Residuals	Std Res.
			Count	%	Count	%		
5-second	Not Interrupted	Incorrect	40,500	20,9%	40,500	20,9%	0,000	0,000
		Correct	4,500	2,3%	4,500	2,3%	0,000	0,000
	Interrupted	Incorrect	47,500	24,5%	47,500	24,5%	0,000	0,000
		Correct	6,500	3,4%	6,500	3,4%	0,000	0,000
15-second	Not Interrupted	Incorrect	38,500	19,8%	38,500	19,8%	0,000	0,000
		Correct	7,500	3,9%	7,500	3,9%	0,000	0,000
	Interrupted	Incorrect	39,500	20,4%	39,500	20,4%	0,000	0,000
		Correct	13,500	7,0%	13,500	7,0%	0,000	0,000

Table 14. Message recall: Cell counts and residuals

Looking at the K-way and higher order effects table on the following page (Table 15), the removal of the one-way effects (the main effects of length, interruption and message recall) and the higher order effects result in statistical significance ($p < 0,001$). The second row ($K = 2$), which removes only the two-way effects (length x interruption, length x recall, interruption x recall) and the three-way effect (length x interruption x recall), is not

statistically significant ($p = 0,240$), as is $K = 3$ ($p = 0,685$), with the removal of the three-way effect.

The ‘K-way effects’ section shows that $K = 1$ yields a significant result ($p < 0,001$), determining that the model would be significantly affected if only the main effects were removed. Removing only the two-way effects, as tested in $K = 2$, the results do not show statistical significance ($p = 0,149$). Similarly, $K = 3$ also produces a non-significant result ($p = 0,685$) after removing the three-way effect.

The tables thus far have presented results indicating a no statistically significant interaction between advertising length, interruption and brand recall ($p = 0,685$), which is relevant to the hypothesis concerning interruption as a moderating variable. Additionally, the two-way effects yield non-significant results ($p = 0,240$), however, one must look at the partial associations to specify exactly which interaction it concerns.

	K	df	Likelihood Ratio		Pearson		Iterations
			Chi-Square	Sig.	Chi-Square	Sig.	
K-way and Higher Order Effects	1	7	108,656	<0,001	96,474	<0,001	0
	2	4	5,495	0,240	5,658	0,226	2
	3	1	0,165	0,685	0,166	0,684	3
K-way Effects	1	3	103,161	<0,001	90,816	<0,001	0
	2	3	5,330	0,149	5,492	0,139	0
	3	1	0,165	0,685	0,166	0,684	0

Table 15. Message recall: K-Way and higher-order effects

If we look at the partial associations in Table 16, length x interruption is not statistically significant ($p = 0,760$), as is interruption x message recall ($p = 0,254$). However, length x message produces a statistically significant result ($p = 0,043$). Therefore, one of the three two-way interactions have statistical significance. Of interest to the hypothesis is the interaction between advertisement length and message recall.

Effect	df	Partial Chi-Square	Sig.	Iterations
Length*Interruption	1	0,093	0,760	2
Length*Message Recall	1	4,080	0,043	2
Interruption*Message Recall	1	1,301	0,254	2
Length	1	0,000	1,000	2
Interruption	1	1,321	0,250	2
Message Recall	1	101,840	<0,001	2

Table 16. Message recall: Partial associations

When analysing the z-scores in the parameter estimates (see Table 17 on the following page), message recall has the largest absolute value ($z = 8,359$), resulting in the strongest effect. Length x message recall has the second highest z-score ($z = 1,790$). The main effects of length and interruption have similar effects, with z-scores of -1,220 and -1,383 respectively.

Effect	Estimate	Std. Error	Z	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
Length*Interruption*Message Recall	-0,044	0,103	-0,429	0,668	-0,246	0,158
Length*Interruption	0,011	0,103	0,105	0,917	-0,191	0,213
Length*Message Recall	0,185	0,103	1,790	0,073	-0,018	0,387
Interruption*Message Recall	0,096	0,103	0,934	0,350	-0,106	0,298
Length	-0,126	0,103	-1,220	0,222	-0,328	0,076
Interruption	-0,143	0,103	-1,383	0,167	-0,345	0,060
Message Recall	0,862	0,103	8,359	0,000	0,660	1,064

Table 17. Message recall: Parameter estimates

After elimination, the final output produces the likelihood ratio statistic based on the new model shown in Table 18. It tests whether the expected frequencies generated by the model are significantly different from the actual frequencies. In that sense, it is favourable for the tests to be non-significant, as they are in this case ($p = 0,591$), indicating a good fit.

Ad Length	Interruption	Message Recall	Observed		Expected		Residuals	Std Res.
			Count	%	Count	%		
5-second	Not Interrupted	Incorrect	40,000	20,6%	43,500	22,4%	-3,500	-0,531
		Correct	4,000	2,1%	5,000	2,6%	-1,000	-0,447
	Interrupted	Incorrect	47,000	24,2%	43,500	22,4%	3,500	0,531
		Correct	6,000	3,1%	5,000	2,6%	1,000	0,447
15-second	Not Interrupted	Incorrect	38,000	19,6%	38,500	19,8%	-0,500	-0,081
		Correct	7,000	3,6%	10,000	5,2%	-3,000	-0,949
	Interrupted	Incorrect	39,000	20,1%	38,500	19,8%	0,500	0,081
		Correct	13,000	6,7%	10,000	5,2%	3,000	0,949

Table 18. Message recall: Cell counts and residuals based on the adjusted model

4.2.1 Interpretation of Message Recall Results

As with brand recall, the results were interpreted using a cross-tabulation with frequencies for each of the categories (Table 19 on the following page), as well as a graphical presentation of the data, as seen in Figure 6 below.

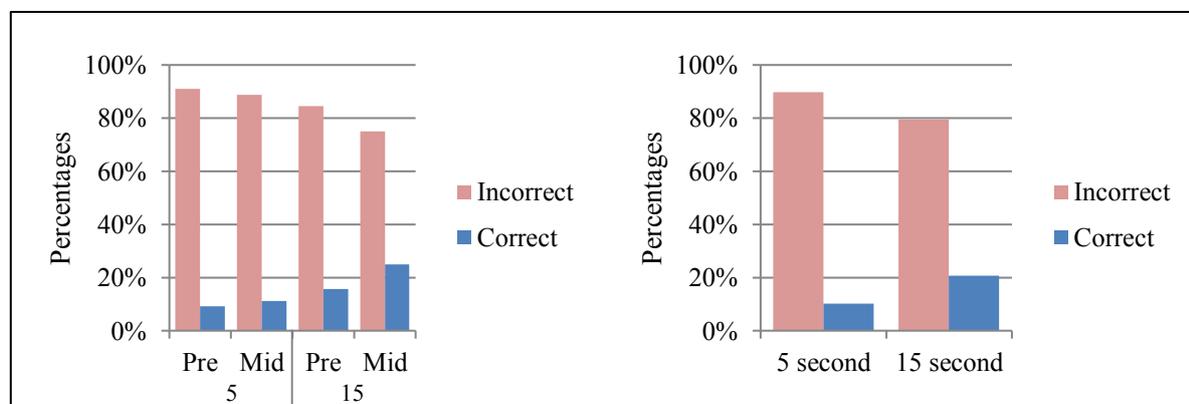


Figure 6. Message recall: Graphical presentation of frequencies

In terms of hypothesis H1b, which tests the relationship between advertisement length and message recall, the right graph in Figure 6 displays the superiority of 15-second advertisements with 15-second advertisements being twice as effective as 5-second advertisements in terms of message recall, with 20,6% vs 10,3% recall respectively. Table 19 displays these exact percentages with respect to each cell.

Ad Length	Interruption	Brand Recall	Count	% within group
5-second	Not Interrupted	Incorrect	40	90,9%
		Correct	4	9,1%
	Interrupted	Incorrect	47	88,7%
		Correct	6	11,3%
	Total	Incorrect	87	89,7%
		Correct	10	10,3%
15- second	Not Interrupted	Incorrect	38	84,4%
		Correct	7	15,6%
	Interrupted	Incorrect	39	75,0%
		Correct	13	25,0%
	Total	Incorrect	77	79,4%
		Correct	20	20,6%

Table 19. Message recall: Cross-tabulation with frequencies

In terms of hypothesis H2b, which tests whether interruption strengthens the positive relationship between advertisement length and message recall, looking at the left graph in Figure 6, one could form an impression that interruption strengthens the relationship between advertisement length and message recall. However, the loglinear analysis resulted in nonsignificant results, causing this interpretation to be void.

In summary, a loglinear analysis has revealed statistically significant relationships only between advertisement length and message recall in respect to hypotheses H1b and H2b. As a result of this conclusion:

→ **We accept the hypothesis H1b and conclude that the length of online video advertisements has a positive effect on message recall.**

→ **We reject the hypothesis H2b and conclude that interruption does not strengthen the positive relationship between online video advertisement length and message recall.**

4.3 Results for Advertisement Likeability

As previous literature suggests, a significant positive relationship exists between advertisement length and advertisement likeability (Newstead & Romaniuk, 2009). Previous studies have also shown that unlike with brand and message recall, interruption inverts the positive relationship between advertisement length and likeability (Goodrich, Schiller & Galletta, 2015; McCoy et al., 2008). Correspondingly, the last two hypotheses test whether such relationships are present for 5-second and 15-second advertisements.

H1c: 15-second online video advertisements generate better advertisement likeability than 5-second online video advertisements.

H2c: Interruption inverts the positive relationship between online video advertisement length and advertisement likeability.

A two-way ANOVA was conducted with the aim of testing the main effect of length on likeability as well as the effect of interruption as a moderating variable. Interruption is expected to cause a weakening effect between the variables of length and likeability, if significant (Goodrich, Schiller & Galletta, 2015). The effect of the moderating variable is illustrated by the interaction effect, which represents the impact of advertisement length depending on the presence of interruption.

The two-way ANOVA initially produces descriptive statistics (Table 20) which in this case suggests that no relationship exists between advertisement length and likeability, with 5-second advertisements and 15-second advertisements producing mean scores of 2,54 and 2,52 respectively. The results for interruption, on the other hand, may cause one to believe that it causes the relationship between advertisement length and likeability to invert as hypothesised. Overall, interrupted advertisements are less liked than non-interrupted advertisements.

Ad Length	Interruption	Mean	Std. Deviation	N
5-second	Not Interrupted	2,59	0,787	44
	Interrupted	2,49	0,724	53
	Total	2,54	0,751	97
15-second	Not Interrupted	2,67	0,674	45
	Interrupted	2,38	0,745	52
	Total	2,52	0,723	97
Total	Not Interrupted	2,63	0,729	89
	Interrupted	2,44	0,733	105
	Total	2,53	0,735	194

Table 20. Advertisement likeability: Descriptive statistics

However, when observing the ‘tests of between subject effects’ table (see Table 21 on the following page), all effects were statistically insignificant, making the latter observations regarding interruption void. The table reveals no main effect for length ($F = 0,020$, $p = 0,887$), indicating no significant difference between advertisements of 5-seconds ($M = 2,54$, $SD = 0,751$), and 15-seconds ($M = 2,52$, $SD = 0,723$), seen in Table 20. Additionally, no main effect for interruption ($F = 3,274$, $p = 0,072$) or the interaction effect ($F = 0,739$, $p = 0,391$). These results were also reflected in the eta squared effect sizes, which were all weak.

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Corrected Model	2.182 ^a	3	0,727	1,352	0,259	0,021
Intercept	1236,304	1	1236,304	2298,651	<0,001	0,924
Length	0,011	1	0,011	0,020	0,887	0,000
Interruption	1,761	1	1,761	3,274	0,072	0,017
Length * Interruption	0,398	1	0,398	0,739	0,391	0,004
Error	102,189	190	0,538			
Total	1342,000	194				
Corrected Total	104,371	193				

Table 21. Advertisement likeability: Test of between-subjects effects

4.3.1 Interpretation of Advertisement Likeability Results

The graphical presentation of the means for each experimental group can be seen in Figure 6. The graph clearly shows that means cluster around the average value of $M = 2,5$. Given the scale of the data, which is from 1 - 5, it is clear that no relationship exists.

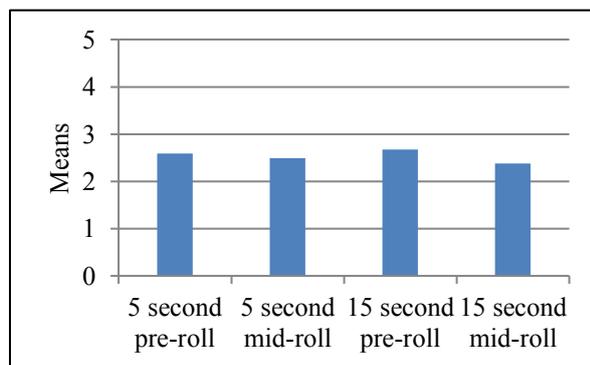


Figure 7. Advertisement likeability: Graphical presentation of means

In summary, the two-way ANOVA has revealed that no statistically significant relationships exist between advertisement length and advertisement likeability, regardless of interruption, in respect to hypotheses H1b and H2b. As a result of this conclusion:

→ We reject hypothesis H1c and conclude that there is no significant relationship between online video advertisement length and advertising likeability.

→ We reject hypothesis H2c and conclude that the interruption does not strengthen the positive relationship between online video advertisement length and advertisement likeability.

4.4 Overview of Tested Hypotheses

This section provides an overview of the above tested hypothesis, in respect to their status. Table 22 on the following page presents the hypotheses in the order they were tested in. As shown below, three hypotheses were accepted in relation to brand recall and message recall.

Hypotheses	Description	Status
H1a	<i>15-second online video advertisements generate better brand recall than 5-second online video advertisements.</i>	Accepted
H2a	<i>Interruption strengthens the positive relationship between online video advertisement length and brand recall.</i>	Accepted
H1b	<i>15-second online video advertisements generate better message recall than 5-second online video advertisements.</i>	Accepted
H2b	<i>Interruption strengthens the positive relationship between online video advertisement length and message recall.</i>	Rejected
H1c	<i>15-second online video advertisements generate better advertisement likeability than 5-second online video advertisements.</i>	Rejected
H2c	<i>Interruption inverts the positive relationship between online video advertisement length and advertisement likeability.</i>	Rejected

Table 22. Tested hypotheses and their status

5 Discussion

Marketers have many options when promoting video advertisements online, in terms of both length and interruption. YouTube offers combinations of the two, with the choice falling in the hands of the marketer. These choices influence the effectiveness of the advertisement, particularly in respect to brand recall, message recall and advertisement likeability. This study aimed to empirically test these effects on 194 experiment participants and was analysed through statistical tests of two-way ANOVA and loglinear analysis. In this chapter, the results will be discussed with respect to previous findings and theory.

5.1 The main effect of advertisement length on advertisement effectiveness

The first question in this study sought to determine whether a relationship exists between advertisement length and effectiveness, specifically brand recall, message recall and advertisement likeability. It was hypothesised that longer, 15-second advertisements would be superior to shorter 5-second versions according to previous findings (Newstead & Romaniuk, 2009). These hypotheses were also in accordance with established knowledge that positively correlates advertisement exposure (length) with memorability (Ha & McCann, 2008) and persuasion (Hovland, 1953).

The results of the current study found that a positive relationship between advertisement length and brand recall existed, as well as between advertisement length and message recall. The most obvious finding to emerge from the analysis was that brand recall was the strongest variable in terms of significance ($p = 0,004$). However, a more insightful result can be observed when looking at Table 13 in the results section. Here, one could see that brand recall rates were weak not only for 5- but also 15-second advertisements, regardless of the significant difference between the two. The table shows that only 14,4% of participants who viewed the 5-second advertisements could recall the brand. However, a more promising rate can be observed for 15-second advertisements, with 30,9% of participants able to recall the brand (Table 13). Message recall was less significant, but remained under the 0,05 confidence level ($p = 0,043$). Message recall provided similar results to brand recall in that recall rates were surprisingly weak – in fact, even more so. Table 19, also located in the results section, shows that only 10,3% of participants who viewed the 5-second advertisements could recall the message. Similarly, only 20,6% of participants who viewed the 15-second advertisements could recall the message. Another unanticipated finding was that no relationship existed for the variable of advertisement likeability whatsoever. Interestingly, Figure 7 in the results section shows that all four experiment groups produced similar likeability scores, at the average score of 2,53 (on a 1-5 scale). This is a curious find and suggests that advertisements are generally disliked, regardless of length. In general, the above findings raise intriguing questions regarding the effectiveness of shorter online video advertisements in respect to all tested variables. Although a significance exists between advertisement length and recall rates, both advertisement lengths produced surprisingly weak results, questioning the benefits of shorter advertisements in general – be it 5-second or 15-second. Such a finding was

unprecedented in respect to previous research. Studies examining shorter advertisements have found them to be relatively effective compared to longer advertisements (Newstead & Romaniuk, 2009; Singh & Cole, 1993).

A strong relationship between advertisement length and advertisement effectiveness in terms of brand recall, message recall and advertisement likeability has been reported in the literature. One of the closest studies to ours was done by Newstead and Romaniuk (2009) who found a significant relationship between advertisement length and the said variables. Our results are mostly in accord with Newstead and Romaniuk's (2009) study, apart from our insignificant likeability results. The results for brand recall also corroborate the ideas of a more recent study by Li and Lo (2015), who suggested that the positive relationship between advertisement length and brand recall continues to exist in the online environment. However, our study differs from previous studies in that our results suggest neither 5-second nor 15-second advertisements to be particularly effective. Although no research has been conducted specifically for 5-second advertisements with respect to brand recall, message recall and advertisement likeability, we can compare our results for 15-second advertisements in this regard. Newstead and Romaniuk's study on television advertisements found that brand recall rates for 15-second advertisements were 42%, compared to our study whereby only 30,9% recalled the brand for the 15-second advertisements. Message recall was found to be at 39% for 15-second television advertisements (Newstead and Romaniuk, 2009), compared to our finding of 20,6%. Furthermore, Newstead and Romaniuk's likeability scores for 15-second advertisements were 3,2 (from a scale of 1-5), while our results fell at 2,52 on the same scale. This difference may be explained by the difference in methods between the studies. For one, Newstead and Romaniuk (2009) conducted a field experiment, utilizing both established and unknown brands which may have skewed the results. However, the weaker results from our side may suggest a difference in cognitive processing in online environments compared to offline environments. This may be further explained by the theory introduced in the earlier sections of the study.

Advertisements rely on the generation of adequate attention and interest, in order to increase memorability and persuasion of the advertisement (Ha & McCann, 2008; Hovland, 1953). Memorability of an advertisement is heightened by exposure time, as is persuasion (Ha & McCann, 2008; Hovland, 1953). This theory relies on the 'total time hypothesis' which assumes a linear relationship between the amount of exposure to the advertisement and its memorability (Pechmann & Stewart, 1988; Rethans, Swasy & Marks, 1986). As referenced in prior sections, the theory is grounded by the knowledge that longer advertisements provide repetition of the brand name and message, and in turn aiding memorability (Fabian, 1986; Mord & Gilson, 1985; Singh & Cole, 1993). Furthermore, persuasion is dependent on the interest towards the advertisement, which of course depends on attention. Such is why the components of the AIDA model do not work in isolation (Gupta, Singh & Sinha, 2016). This point may be of aid when explaining the difference in results between our study and that of Newstead and Romaniuk (2009). For one, television advertisements were given more attention, causing recall rates to strengthen. Resultantly, likeability scores were higher, as a direct function of attention. In other words, because our online video advertisements received less attention, the interest towards the advertisements suffered as a result. This raises the question of why online advertisements attracted less attention compared to offline advertisements. To answer this question, we can once again fall back on the theory surrounding cognitive processing. As was found by Ha and McCann (2008) and Yoon, Choi and Sujin (2011), people are more goal-oriented when watching videos online than offline. One specific theory concerning this is named goal pursuit, in which people actively seek

information that is relevant to them (Laran, Janiszewski & Salerno, 2016). In this sense, “being aware of a goal enables people to ... identify attributes that are relevant to [their] goal pursuit” (Laran, Janiszewski & Salerno, 2016, p.442). As a direct result, people tend to ignore information irrelevant to them in the quest for more relevant information. Unsurprisingly, advertisements receive less attention and interest in online environments and is what could be used to explain our mediocre recall and likeability results.

Overall, our findings provide empirical evidence on the effects of alternative advertisement lengths on brand recall and message recall and advertisement likeability in the online environment. Results were mostly aligned with previous research, besides results for advertisement likeability. However, due to the purely quantitative methods used, this study could not shed further analysis into why this finding regarding advertisement likeability occurred. This may very well be another important issue for future research on the topic. What is empirically distinct in our results is that advertisements 15-second or less produce little benefits in increasing overall attention and interest surrounding a brand. Although a significance exists, it is important to understand this implication because advertisers should not rely on length alone to increase effectiveness.

5.2 The interaction of advertisement length, interruption and advertisement effectiveness

Our second research question aimed at understanding the role of interruption as a moderating variable in respect to the relationship between advertisement length and effectiveness. Hypotheses presented the assumption that interruption (1) strengthens the positive relationship between advertisement length and recall (both brand and message) and (2) inverts the positive relationship between advertisement length and advertisement likeability. It was hypothesised that 15-second interruption-based advertisements would be superior to their 5-second counterparts in respect to brand and message recall; while the opposite is hypothesised for the variable of likeability.

The results of the current study found that interruption positively moderated the relationship between advertisement length and brand recall. Again, the most obvious finding to emerge from the analysis was that brand recall was the strongest dependent variable in terms of significance ($p = 0,041$), in fact the only one with significance. An important finding to arise from the analysis was seen in Figure 5 in the results section. When observing the results graphically, one can see that 15-second interrupted advertisements skew the results significantly. In fact, the results for 5-second advertisements show little increase when interrupted. Brand recall for 5-second advertisements rises a mere 1,5% when interruption-based, as seen in Table 13. However, a significant rise in brand recall for 15-second advertisements occurs with the presence of interruption. Brand recall rises 32,9% in this case (Table 13). In terms of message recall and advertisement likeability, no significant results suggest interruption moderates its relationship with advertisement length. As stated before, likeability scores were constant throughout all experiment groups, as seen in Figure 7. Message recall was weak throughout, with no experiment groups topping 25% recall, as seen in Table 19. 15-second interrupted advertisements produced the highest score, at 25%. Again, once observing the split between interrupted and non-interrupted advertisements in Figure 5, it becomes apparent that 15-second interrupted advertisements skew results. All other groups

fall below 15% message recall (Table 19). These findings were unanticipated; insignificant findings for message recall and likeability raise questions. One could question whether 5-second advertisements are just too short to produce adequate message recall. Moreover, although brand recall is said to be the most important variable for measuring advertisement effectiveness, message recall and advertisement likeability remain relatively important and should be at the forefront to any advertising effort (Newstead & Romaniuk, 2009). Therefore, a second question arises whether 5-second advertisements adequately balance all three variables. According to the above results, 5-second advertisements were relatively ineffective at providing adequate brand recall, message recall and advertisement likeability, regardless of interruption.

Although the effect of interruption as a moderating variable has received a lack of attention when it comes to research, its main effect has been tested by Li and Lo (2015) and can be somewhat compared in this regard. The authors found a significant difference in brand recall rates when comparing pre-roll and mid-roll advertisements, with the latter receiving superior recall results (Li & Lo, 2015). Our results are in accord with Li and Lo's study regarding this variable. However, a lack of similar research exists for the variables of message recall and likeability, resulting in a lack of comparison. With respect to these variables, theory may aid in interpreting our results.

These findings relate to the way we process information online. In the environment of YouTube, watching informative videos involves actively processing information, more so than in television (Ha & McCann, 2008). When disrupted, viewers adopt a similar level of cognitive activity for the advertisement as with the original YouTube video (Li, Edwards & Lee, 2002). This is known as the spill over effect (Sahni, 2016). As such, learning of the advertisement content is heightened when the advertisement is interruption-based because of this spill over effect (Li, Edwards & Lee, 2002; Sahni, 2016). This superior mental engagement provides the assumption that interruption moderates the relationship between advertisement length and effectiveness (Li, Edwards & Lee, 2002). However, an important distinction should be made for the variable of likeability. The latter concerns one's attitude towards the advertisement, which, intuitively, does not perform well in interruption-based environments. Goal pursuit theory suggests that goal-orientated people who are interrupted will feel increased dislike towards advertisements that span longer durations (Goodrich, Schiller & Galletta, 2015). In this respect, interruption is theoretically assumed to invert the relationship between length and likeability. Although brand recall results are validated by theory, the insignificant results for message recall and likeability are unaccounted for. One explanation for the insignificant findings in message recall and advertisement likeability could be the fact that people are much more goal-orientated in online environments, which works to their disadvantage (Ha & McCann, 2008; Yoon, Choi & Sujin, 2011). Considering this, it is much more likely for people to filter out unnecessary information on the quest for more relevant information, as well as be more irritated when disrupted (McCoy et al., 2008). This may raise the point that people choose not to pay attention to advertisements at all, causing the low message recall and likeability scores in general.

Overall, our findings for the interaction effect provide empirical evidence on the effectiveness of short advertisements in the online environment. What is empirically distinct in these results is that 15-second interrupted advertisements are far more superior in terms of brand recall than any other experiment group. Resultantly, 5-second advertisements provide little benefits in increasing overall attention and interest surrounding a brand, regardless of interruption.

6 Conclusion

6.1 Research Aims and Objectives

As online advertising becomes the dominating type of promotion in the 21st century (Statista, 2017a), understanding how each of its forms work is imperative to the success of online advertising in general. 5-second video advertisements are among the newest forms of online advertising seen today. They represent the direction that the advertising scene is taking and reflect the fast-paced audience of the online environment (Microsoft, 2015). This study aimed to shed light on their effectiveness by examining how they performed with respect to memory and attitude towards the advertisement. To gain better insight, these ultra-short advertisements were also tested in an interrupting setting that, according to previous studies, supposedly improves their effectiveness (Gupta, Singh & Sinha, 2016). In order to see where 5-second advertisements stood in comparison to other forms of advertisements, they were compared to their longer 15-second versions, a popular format that has been tested and used in the advertising industry (Singh, Linville & Sukhdial, 1995).

As predicted, 5-second video advertisements did not perform well regarding memorability. Even when used in an interrupting setting, a condition that was predicted to improve their performance, they did not manage to generate significantly better results. They did worse than their longer counterparts, proving length to be a deciding factor. When it came down to the attitudes towards the advertisement, the 5-second advertisements did not perform well either. However, this time the ultra-short advertisements were on par with 15-second advertisements, suggesting 5-second advertisements to be as effective as 15-second advertisements in this regard.

6.2 Theoretical and Practical Implications

6.2.1 Theoretical Implications

Our study holds several theoretical implications regarding cognitive reactions to ultra-short advertisements, with respect to attention and interest. First, our results suggest that advertisement length is important for online video advertising, with respect to brand recall and message recall, as found for television advertisements (Newstead & Romaniuk, 2009). However, a more important contribution is that 5-seconds are relatively inadequate at generating brand and message recall, while the longer, 15-second versions of the advertisements generated significantly better results. This would suggest that adequate learning relies on lengths greater than 5-seconds. This contributes specifically to the theory of the 'total time hypothesis' (Pechmann & Stewart, 1988; Rethans, Swasy & Marks, 1986), whereby we propose a limit in advertisement length exists in order to effectively learn advertisement content. Secondly, this study found that interest towards the advertisement did

not alter in terms of likeability scores. This has theoretical implications in that advertisement length may not contribute to likeability below a certain length. However, this last finding may be the result of extraneous factors, which would require further investigation.

Additionally, few studies have examined the interaction effect of advertisement length, interruption and advertisement effectiveness, with a significant gap in regarding message recall and advertisement likeability. Our study contributes in this regard, in understanding whether theory and practice align. It was found that interruption significantly moderates the relationship between advertisement length and brand recall. However, no such relationship existed between message recall and likeability. This suggests that interruption has a greater effect on brand recall than message recall or likeability.

6.2.2 Practical Implications

Online advertising accounts for a third of worldwide advertising spending (Statista, 2017a) and a strong trend is rising towards shorter video advertisements (Jeong, 2015; Singh, Linville & Sukhdial, 1995). Therefore, it should not come as a surprise that this study forms several practical implications as well. The study demonstrated how effective an ultra-short online video advertisement can be. Moreover, advertisers can gain insights into how these advertisements perform against another widely used format – 15-second advertisements (Jeong, 2015; Singh, Linville & Sukhdial, 1995).

Since effectiveness was measured through three isolated variables, the findings of the study can be utilised by advertisers with different needs, whether wanting to study brand recall, message recall or advertisement likeability. For advertisers considering using 5-second advertisements to increase awareness, brand or message recall should be of more concern. For example, when dealing with unknown brands, short advertisements do not seem to be the best way to go, regardless of whether they are interrupting or not. Advertisers concerned more with adjusting consumer interest to their product or brand should focus on the variable of advertisement likeability. Looking at it in this case, advertisers seeking means to improve their brand image could make a conclusion that using online video advertising, in general, may not be right strategy.

Finally, even though the study primarily tested YouTube advertisements, it does not mean that its findings are reserved for YouTube advertisements only. As mentioned before, shorter video advertisements are growing in popularity, resulting in them being adopted by various websites, including the world's largest social media network Facebook (Facebook, 2017, Statista, 2017a). As such, a rise in the popularity of 5-second advertisements calls for additional research on the topic, for which this study hopes to pioneer.

6.3 Limitations and Future Research

Throughout the duration of the study, we engaged in various efforts to ensure valid and reliable research through the use of a representative sample, appropriate measures and accurate statistical analysis. Nevertheless, the study came with its fair share of limitations despite most them being addressed and adjusted for. The reason why some of these limitations are pointed out or repeated in this last part of the paper is because we see them as promising for laying grounds for future research.

Firstly, the use of convenience sampling strongly influenced the resulting demographics of the sample, being mostly younger Lithuanians aged between 18-24 years. Future research could rather adopt a probability sampling technique that would generate a larger and more diverse sample size. This in turn would provide insights into cross-cultural effects and would help in comprehending how the tested advertisements affect internet users on a broader scale.

Secondly, the study exclusively tested the effectiveness of advertisements for an unknown brand and without celebrity endorsement. Therefore, the study isolated its respondents from such influences that would alter how the advertisements is processed (Dens & De Pelsmacker, 2010). In future investigations, it would be beneficial to test the effectiveness of well-known brands and examine how these advertisements compare to lesser known brands. The combination of findings would presumably give a significantly deeper understanding into the practical applicability of the tested advertisement types.

As mentioned before, once selecting both the original YouTube video and the advertisement, congruity was considered. This concerned matching the content of the advertisement with original YouTube video, and is a specific feature that YouTube advertisers can use (Google, 2017b). However, there are many more targeting features that the study had not adjusted for. For example, YouTube allows advertisers to accurately narrow down their scope in respect to the location, language, demographics and interests of their target audience (Google, 2017b). Thus, further investigation and experimentation using these features is recommended, as it may increase the effectiveness of the tested advertisement.

Finally, the study was restricted in using only one original YouTube video between each of the experimental groups. The length of the video spanned a duration of 2 minutes, and was selected in order to satisfy the minimum length requirements for showing advertisements within videos (Google, 2017a), as well as to balance the short attention spans of respondents. Even though the length proved to be fitting, there is abundant room for further research in determining how effective the tested advertisements would have been in combination with longer videos. Therefore, a similar study using original videos of different lengths would provide a broader understanding of online video advertising and would lay grounds for understanding which combinations of the two videos work best.

7 References

- Albers-Miller, N. D. & Stafford, M. (1999). An International Analysis of Emotional and Rational Appeals in Services vs Goods Advertising, *Journal of Consumer Marketing*, vol. 16, no. 1, pp.42–57.
- Allan, D. (2007). Comparative Effectiveness of 30-versus 60-Second Radio Commercials on Recall and Rate, *Journal of Radio Studies*, vol. 14, no. 2, pp.165–177.
- Bettman, J. R. & Park, W. C. (1980). Effects of Prior Knowledge and Experience and Phase of the Choice Process on Consumer Decision Processes: A Protocol Analysis, *Journal of Consumer Research*, vol. 7, no. 3, pp.234–248.
- Biel, A. L. & Bridgwater, C. A. (1990). Attributes of Likable Television Commercials, *Journal of Advertising Research*, vol. 30, no. 3, pp.38–44.
- Bryman, A. & Bell, E. (2011). *Business Research Methods*, Oxford: Oxford University Press.
- Burns, R. B. & Burns, R. A. (2008). *Business Research Methods and Statistics Using SPSS*, London: Sage Publications.
- BuzzFeed. (2017). BuzzFeed Advertise: Content Worth Sharing, Available Online: <https://www.buzzfeed.com/advertise> [Accessed 19 April 2017].
- Chan, J. C. F., Jiang, Z. & Tan, B. C. Y. (2010). Understanding Online Interruption-Based Advertising: Impacts of Exposure Timing, Advertising Intent, and Brand Image, *IEEE Transactions on Engineering Management*, vol. 57, no. 3, pp.365–379.
- Chatterjee, P. (2008). Are Unclicked Ads Wasted? Enduring Effects of Banner and Pop-Up Ad Exposures On Brand Memory and Attitudes, *Journal of Electronic Commerce Research*, vol. 9, no. 1, pp.51–61.
- ComScore. (2011). The State of Online Video - comScore, Inc, Available Online: http://www.comscore.com/Insights/Presentations-and-Whitepapers/2011/The-State-of-Online-Video?cs_edgescape_cc=SE [Accessed 3 May 2017].
- ComScore. (2016). What Millennials' YouTube Usage Tells Us about the Future of Video Viewership, Available Online: <https://www.comscore.com/Insights/Blog/What-Millennials-YouTube-Usage-Tells-Us-about-the-Future-of-Video-Viewership> [Accessed 3 May 2017].
- Danaher, P. J. & Mullarkey, G. W. (2003). Factors Affecting Online Advertising Recall: A Study of Students, *Journal of Advertising Research*, vol. 43, no. 3, pp.252–267.
- Dehghani, M., Niaki, M. K., Ramezani, I. & Sali, R. (2016). Evaluating the Influence of YouTube Advertising for Attraction of Young Customers, *Computers in Human Behavior*, vol. 59, pp.165–172.
- Dens, N. & De Pelsmacker, P. (2010). Advertising for Extensions: Moderating Effects of Extension Type, Advertising Strategy, and Product Category Involvement on Extension Evaluation, *Marketing Letters*, vol. 21, no. 2, pp.175–189.
- Easterby-Smith, M., Thorpe, R. & Jackson, P. R. (2015). *Management and Business Research*, London: Sage Publications.

- Elsen, M., Pieters, R. & Wedel, M. (2016). Thin Slice Impressions: How Advertising Evaluation Depends on Exposure Duration, *Journal of Marketing Research*, vol. 53, no. 4, pp.563–579.
- Fabian, G. S. (1986). 15-Second Commercials: The Inevitable Evolution, *Journal of Advertising Research*, vol. 26, no. 4, pp.3–5.
- Facebook. (2017). In-Stream Video: Facebook Help Center, Available Online: <https://www.facebook.com/business/help/257808937971866> [Accessed 20 May 2017].
- Field, A. (2009). *Discovering Statistics Using SPSS*, London: Sage Publications.
- Ghauri, P. & Grønhaug, K. (2005). *Research Methods in Business Studies: A Practical Guide*, London: Pearson Education Limited.
- Goldstein, D. G., McAfee, R. P. & Suri, S. (2011). The Effects of Exposure Time on Memory of Display Advertisements, in *EC: Electronic Commerce*, 2011.
- Goodrich, K., Schiller, S. Z. & Galletta, D. (2015). Consumer Reactions to Intrusiveness of Online-Video Advertisements Do Length, Informativeness, and Humor Help (or Hinder) Marketing Outcomes?, *Journal of Advertising Research*, vol. 55, no. 1, pp.37–50.
- Google. (2016). Top Four Content Categories on YouTube – Think with Google, Available Online: <https://www.thinkwithgoogle.com/data-gallery/detail/top-content-categories-youtube/> [Accessed 8 April 2017].
- Google. (2017a). How Video Ads Work, Available Online: https://support.google.com/displayspecs/answer/6244557?hl=en&ref_topic=6244532 [Accessed 18 May 2017].
- Google. (2017b). AdWords Help: About Targeting for Video Campaigns, Available Online: https://support.google.com/adwords/answer/2454017?hl=en&ref_topic=3119140 [Accessed 3 May 2017].
- Google. (2017c). Advertising Policies Help: Alcohol, Available Online: <https://support.google.com/adwordspolicy/answer/6012382?hl=en> [Accessed 3 May 2017].
- Greene, W. F. (1991). Observations: What Drives Commercial Liking? An Exploration of Entertainment vs. Communication, *Journal of Advertising Research*, vol. 32, no. 2, pp.65–68.
- Gupta, H., Singh, S. & Sinha, P. (2016). Multimedia Tool as a Predictor for Social Media Advertising- a YouTube Way, *Multimedia Tools and Applications*, no. 1, pp.1–12.
- Ha, L. & McCann, K. (2008). An Integrated Model of Advertising Clutter in Offline and Online Media, *International Journal of Advertising*, vol. 27, no. 4, pp.569–592.
- Haley, R. I. & Baldinger, A. L. (2000). The Arf Copy Research Validity Project, *Journal of Advertising Research*, vol. 40, no. 6, pp.114–135.
- Heffernan, V. (2009). We Interrupt This Program, *The New York Times Magazine*, 2 January, Available Online: <http://www.nytimes.com/2009/01/04/magazine/04wwln-medium-t.html> [Accessed 1 April 2017].
- Hollis, N. (1995). Like It Or Not, *Journal of Advertising Research*, vol. 35, no. 5, pp.7–16.
- Hovland, C. I. (1953). *Communication and Persuasion : Psychological Studies of Opinion*

Change, New Haven: Yale University Press.

- Jankowski, J., Kazienko, P., Wątróbski, J., Lewandowska, A., Ziemia, P. & Ziolo, M. (2016). Fuzzy Multi-Objective Modeling of Effectiveness and User Experience in Online Advertising, *Expert Systems with Applications*, vol. 65, no. 1, pp.315–331.
- Jeong, Y. (2015). The Impact of the Length of Preceding and Succeeding Ads on Television Advertising Effectiveness, *Journal of Marketing Communications*, pp.1–15.
- Jeong, Y., Sanders, M. & Zhao, X. (2011). Bridging the Gap between Time and Space: Examining the Impact of Commercial Length and Frequency on Advertising Effectiveness, *Journal of Marketing Communications*, vol. 17, no. 4, pp.263–279.
- Johnson, E. J. & Russo, E. J. (1984). Product Familiarity and Learning New Information, *Journal of Consumer Research*, vol. 11, no. 1, pp.542–550.
- Kapferer, J.-N. (1997). *Strategic Brand Management : Creating and Sustaining Brand Equity Long Term*, London: Kogan Page.
- Keller, K. L. (2008). *Strategic Brand Management : Building, Measuring, and Managing Brand Equity*, New Jersey: Pearson.
- Krugman, H. E. (1983). Television Program Interest and Commercial Interruption, *Journal of Advertising Research*, vol. 23, no. 1, pp.21–23.
- Laczniak, R. N. (2015). The Journal of Advertising and the Development of Advertising Theory: Reflections and Directions for Future Research, *Journal of Advertising*, vol. 44, no. 4, pp.429–433.
- Laran, J., Janiszewski, C. & Salerno, A. (2016). Exploring the Differences Between Conscious and Unconscious Goal Pursuit, *Journal of Marketing Research*, vol. 53, no. 3, pp.442–458.
- Li, H., Edwards, S. M. & Lee, J.-H. (2002). Measuring the Intrusiveness of Advertisements: Scale Development and Validation, *Journal of Advertising*, vol. 31, no. 2, pp.37–47.
- Li, H. & Lo, H.-Y. (2015). Do You Recognize Its Brand? The Effectiveness of Online In-Stream Video Advertisements, *Journal of Advertising*, vol. 44, no. 3, pp.208–218.
- Lloyd, D. W. & Clancy, K. J. (1991). CPMs versus CPMIs: Implications for Media Planning, *Journal of Advertising Research*, vol. 31, no. 4, pp.34–44.
- Lundqvist, A., Liljander, V., Gummerus, J. & van Riel, A. (2012). The Impact of Storytelling on the Consumer Brand Experience: The Case of a Firm-Originated Story, *Journal of Brand Management*, vol. 20, no. 4, pp.283–297.
- Malhotra, N. K. (2010). *Marketing Research: An Applied Orientation*, 6th edn, Upper Saddle River: Prentice Hall.
- Martin-Santana, J. D., Reinares-Lara, P., Rey, U., Carlos, J., Universidad, E. R.-L. & Carlos, R. J. (2016). Spot Length and Unaided Recall in Television, *Journal of Advertising Research*, vol. 53, no. 3, pp.274–288.
- McCarthy, J. E. (1960). *Basic Marketing: A Managerial Approach*, Homewood: Irwin.
- McCoy, S., Everard, A., Polak, P. & Galetta, D. F. (2008). An Experimental Study of Antecedents and Consequences of Online Ad Intrusiveness, *Journal of Human-Computer Interaction*, vol. 24, no. 7, pp.672–699.

- Microsoft. (2015). Attention Spans, Available Online:
<https://advertising.microsoft.com/en/.../microsoft-attention-spans-research-report.pdf>.
- Moorman, M., Neijens, P. C. & Smit, E. G. (2002). The Effects of Magazine-Induced Psychological Responses and Thematic Congruence on Memory and Attitude Toward the Ad in a Real-Life Setting, *Journal of Advertising*, vol. 31, no. 4, pp.27–40.
- Mord, M. S. & Gilson, E. (1985). Shorter Units: Risk-Responsibility-Reward, *Journal of Advertising Research*, vol. 25, no. 4, pp.9–19.
- Newell, S. J. & Henderson, K. V. (1998). Super Bowl Advertising: Testing the Importance of Advertisement Frequency, Length and Placement on Recall, *Journal of Marketing Communications*, vol. 4, no. 4, pp.237–248.
- Newstead, K. & Romaniuk, J. (2009). Cost per Second: The Relative Effectiveness of 15- and 30-Second Television Advertisements, *Journal of Advertising Research*, vol. 50, no. 1, pp.68–76.
- Pechmann, C. & Stewart, D. W. (1988). Advertising Repetition: A Critical Review of Wearin and Wearout, *Current Issues & Research in Advertising*, vol. 11, no. 2, pp.285–329.
- Petty, R. E., Cacioppo, J. T. & Schumann, D. (1983). Central and Peripheral Routes to Advertising Effectiveness: The Moderating ...: Lund University Libraries, *Journal of Consumer Research*, vol. 10, no. 2, pp.135–146.
- Rethans, A. J., Swasy, J. L. & Marks, L. J. (1986). Effects of Television Commercial Repetition, Receiver Knowledge, and Commercial Length: A Test of Two-Factor Model, *Journal of Marketing Research (JMR)*, vol. 23, no. 1, pp.50–61.
- Richards, J. I. & Curran, C. M. (2002). Oracles on ‘Advertising’: Searching for a Definition, *Journal of Advertising*, vol. 31, no. 2, pp.63–77.
- Saunders, M., Lewis, P. & Thornhill, A. (2007). *Research Methods for Business Students*, Harlow: Pearson.
- Singh, S. N. & Cole, C. A. (1993). The Effects of Length, Content, and Repetition on Television Commercial Effectiveness, *Journal of Marketing Research*, vol. 30, no. 1, pp.91–104.
- Singh, S. N., Linville, D. & Sukhdial, A. (1995). Enhancing the Efficacy of Split Thirty-Second Television Commercials: An Encoding Variability Application, *Journal of Advertising*, vol. 3, no. 24, pp.13–23.
- Smit, E. G., van Meurs, L. & Neijens, P. C. (2006). Effects of Advertising Likeability: A 10-Year Perspective, *Journal of Advertising Research*, vol. 46, no. 1, pp.73–83.
- Speck, P. S. & Elliott, M. T. (1997). Predictors of Advertising Avoidance in Print and Broadcast Media, *Journal of Advertising*, vol. 26, no. 3, pp.61–76.
- Statista. (2017a). Global Ad Spend Distribution by Medium 2016, Available Online:
<https://www.statista.com/statistics/376260/global-ad-spend-distribution-by-medium/>
 [Accessed 25 March 2017].
- Statista. (2017b). Global Social Media Ranking 2017 | Statistic, Available Online:
<https://www.statista.com/statistics/272014/global-social-networks-ranked-by-number-of-users/>
 [Accessed 18 April 2017].
- Strong, E. K. (1925). *The Psychology of Selling and Advertising*, New York: McGraw-Hill.

Wilson Van Voorhis, C. R. & Morgan, B. L. (2007). Understanding Power and Rules of Thumb for Determining Sample Sizes, *Tutorials in Quantitative Methods for Psychology*, vol. 3, no. 2, pp.43–50.

Yoon, S., Choi, Y. K. & Sujin, S. (2011). When Intrusive Can Be Likable, *Journal of Advertising*, vol. 40, no. 2, pp.63–75.

YouTube. (2017a). YouTube Advertising Formats, Available Online:
https://support.google.com/youtube/answer/2467968?visit_id=1-636261241123213121-3349237616&rd=1 [Accessed 26 March 2017].

YouTube. (2017b). Statistics, Available Online:
<https://www.youtube.com/yt/press/statistics.html> [Accessed 6 April 2017].

Appendix A: A correlation matrix

Control Variables		Brand Recall	Message Recall	Ad Likeability
Gender	Pearson Correlation	-0,020	0,108	-0,045
	Sig. (2-tailed)	0,781	0,135	0,531
Age	Pearson Correlation	-0,037	-0,026	-0,135
	Sig. (2-tailed)	0,609	0,720	0,060
Product Familiarity	Pearson Correlation	0,091	0,023	0,114
	Sig. (2-tailed)	0,205	0,752	0,113
Nationality	Pearson Correlation	-0,015	0,028	0,043
	Sig. (2-tailed)	0,840	0,696	0,547

Appendix B: Testing Outliers for ANOVA

