The conscious and non-conscious effects of context in advertising

Peter Andersson    Jesper Ring Jakobsen

Supervisors:
Seidi Suurmets & Tore Kristiansen

September 1st
STU: 222.101
Acknowledgments

We would like to express our deepest gratitude to our main supervisor, Seidi Suurmets.

Seidi has been a crucial part of the project and has been providing invaluable advice and counseling throughout the entire process. Her dedication, engagement and involvement has been far greater than we could have ever asked for, and this thesis would not have been possible without her. We sincerely appreciate everything Seidi has contributed with. Thank you.

Furthermore, we would also like to express our gratitude to Tore Kristiansen, Gorm Gabrielsen and Dalia Bagdziunaite that have all given valuable advice and counseling in regards to theory selection, statistical analysis and experimental considerations.
Executive summary

In today’s fast moving society, the amount of commercial and noncommercial messages consumers can be expected to encounter on a daily basis is enormous. One of the most salient and omnipresent platforms for market communications is advertising. In the face of a new marketing landscape, and with the emergence of an interesting range of new market research techniques, advertising is facing some interesting challenges. One such challenge is to better understand how the context of advertising affects the single ad - a question that in many traditional models of marketing communication is neglected as ads are often analyzed in isolation.

Leaning on a body of theory from marketing communication combined with more recent theories from social psychology and neuromarketing, some contradictory assumptions about consumer and the mechanisms behind advertising is discussed. By using both biometric tools and self-report measures, two experiments were conducted to gain insight into the complex matter of contextual effects. More specifically, consumers conscious and non-conscious reactions to different advertisements placed in a specific context were investigated. The main method of data acquisition has been through pupil size measurement via eye-tracking and related rating tasks in a laboratory setting.

The findings from the experiments presented suggests that context does matter. Automatic, non-conscious reactions such as arousal were, for the majority of investigated advertisements, affected by the context in which they were placed. Conscious evaluations were also affected, however, to a much smaller degree. The findings also indicate that the traditional way of conducting market research might be inadequate for gaining truly meaningful data. The implications, both in an academic and a more market oriented perspective, are discussed. Finally, a revised framework of the value based model of choice, incorporating the influence of external stimuli (such as context) is suggested.
Table of Content

1 General introduction .................................................................................................................. 6
   1.1 Research Question ............................................................................................................. 8
   1.2 Reader’s guide .................................................................................................................... 9
   1.3 Epistemology .................................................................................................................... 10
   1.4 Delimitations .................................................................................................................... 13
   1.5 Concept definition ............................................................................................................. 14
       1.5.1 A brief history of definitions for “Emotions & Feelings” ............................................. 14
       1.5.2 Modern definitions of “Emotions and Feelings” ......................................................... 16

2 Traditional marketing theory and research .............................................................................. 19
   2.1 Traditional perspectives and research of consumer behavior ......................................... 19
       2.1.1 Buyer behaviour ......................................................................................................... 19
       2.1.2 Consumer behavior ................................................................................................... 20
       2.1.3 Consumer research .................................................................................................... 21
       2.1.4 Consumption studies ................................................................................................. 21
       2.1.5 Perspective of this paper ............................................................................................ 22
   2.2 Traditional Advertising & Marketing communication models .......................................... 23
       2.2.1 Hierarchy-of-effects models ....................................................................................... 23
       2.2.2 AIDA & FCB ............................................................................................................... 24
       2.2.3 Elaboration likelihood model ..................................................................................... 26

3 Modern views on consumer behavior .................................................................................... 28
   3.1 Contributions from social psychology .............................................................................. 29
   3.2 Priming ................................................................................................................................ 30
       3.2.1 Associative priming ..................................................................................................... 32
       3.2.2 Motivational priming .................................................................................................. 33
   3.3 Social psychology models ................................................................................................. 34
       3.3.1 Emotional congruence ............................................................................................... 34
       3.3.2 Affect Infusion model ............................................................................................... 36
       3.3.3 Affect-as-information ............................................................................................... 39

4 The role of Neuroscience ..................................................................................................... 41
4.1 Neuromarketing perspectives on consumer decision-making ........................................42
4.1.1 Representation & Attention..................................................................................44
4.1.2 Predicted value ......................................................................................................46
4.1.3 Experienced value, remembered value and learning ..............................................46
4.2 The role of arousal......................................................................................................47
4.2.1 Measuring Arousal ...............................................................................................49
4.3 The role of valence ....................................................................................................53
4.3.1 Measuring valence ...............................................................................................54
4.4 Relationship between arousal and valence...............................................................57
5 Theoretical summary ....................................................................................................59
6 Research design ............................................................................................................60
6.1 Hypothesis definition .................................................................................................62
6.1.1 Correlation between pupil size and subjective valence score ................................62
6.1.2 Effects of ad-viewing in a context rather than isolated .........................................62
6.1.3 The spillover effects of differently valenced images – sequential effects ............63
6.2 Operational definition of variables ..........................................................................64
6.2.1 Arousal ................................................................................................................64
6.2.2 Valence ...............................................................................................................66
7 Group 1 - Experimental design and rationale ...............................................................68
7.1.1 Manipulation of ads .............................................................................................70
7.1.2 Subject sample .....................................................................................................72
7.1.3 Research equipment, software & location ............................................................73
7.1.4 Other Methodological considerations ..................................................................73
7.2 Group 1 - Results .....................................................................................................76
7.2.1 Hypothesis H1a ....................................................................................................77
7.2.2 Hypothesis H1b ....................................................................................................78
7. Selection of ads for group 2 ......................................................................................79
8 Group 2 - Experimental design and rationale ...............................................................82
8.1.1 Manipulation of stimuli .......................................................................................84
8.1.2 Subject sample ....................................................................................................84
8.1.3 Research equipment, software & location ............................................................85
8.1.4 Other methodological considerations ..................................................................85
1 General introduction

In today's society, most consumers will be exposed to an enormous amount of advertising and marketing. This number has by some authors been estimated to total somewhere around 500 and 5000 messages a day (Bovee & Arens, 1995). Due to the enormous popularity of advertising as a marketing tool, the history of research in the field is also rich and continuously growing. The studies of advertisement can be found within many different research areas, however most traditional marketing and advertising models seem to rely on a set of assumptions that in recent time has been challenged.

Traditionally, it has been argued that the consumer is mostly a rational and conscious being (Genco et al., 2013). According to traditional theories of consumer decision making, consumers are “governed by rational and conscious maximization of expected utility, rather than by emotional and nonconscious processes” (Hubert & Kenning, 2008). Because of progress within fields such as consumer behavior studies, social psychology and neuroscience, a growing, and arguably healthy, awareness and skepticism is shown towards this assumption.

Extensive research has in recent times proved that decisions, in fact, are not primarily based upon rational and cognitive processing of information (Damasio, 1994; Pham 1998; Schwartz 2010; Genco et al, 2013). Instead, the traditional cognitive based views on consumers have to some extent been dethroned by theories from fields such as consumer psychology with the demonstration of decision-making as primarily being rooted in nonconscious processes (Kahneman, 2002; Chartrand and Fitzsimons, 2011).

A related implication of this insight is then also the problem of assuming that consumers are consciously able to report their motives, thoughts or even actions. With the technological advances, neuroscientific studies, amongst others, has helped illustrate how these assumptions are often faulty and that people are often relatively unaware of their own motivations and cannot report themselves what drives them and their behavior. If this is the case, then it would seem problematic
that the vast majority of traditional consumer research relies on self-reporting through interviews, focus groups, surveys and other similar techniques.

Studying how the brain works and its underlying mechanisms is an interesting and fascinating pursuit. It has become an important source of insight for many, different fields - amongst them, marketing, because it aids marketers to better tap into the mind of consumers by providing a better understanding of different steps in consumption, decision and purchase behavior.

Not surprisingly, the academic disciplines of neuroscience and neuromarketing has evolved considerably within the last few decades as proven by the sheer number of publications within marketing and consumer behavior journals as well as illustrated by general growth in Google search entries as well as in the number of registered neuromarketing companies (Plassmann, Ramsøy, and Milosavljevic, 2012).

Having established the extensive usage of ads and the enormous exposure of ads that consumers are subject to, the faulty assumptions many traditional marketing theories have and the new possibilities of neuroscience, interesting discussions about the dynamics of advertising emerge.

Since advertising has taken a central role in today’s society, it is extremely rare that consumers are exposed to one single ad in isolation. Whether it be a TV-commercial, webpage banner, magazine ad, etc. almost all form of advertising happens in a context and in conjunction with other competing ads and commercial or noncommercial messages. This fact raises the question of how the context and sequence of messages influences consumers’ perception and judgement of the ads they see.

As established, many traditional theories assume consumers to be rational, however, another critical aspect that is related to the assumptions about how the consumer is believed to behave, is that advertisements and their effectiveness is often created, reviewed and examined in isolation. This is a natural consequence of the assumptions related to the rational consumers, because if the consumer is truly rational, they will be able to sort out what information is relevant for evaluating the ad, product or service at hand. The noise and distortion from competing messages should thus not matter to the truly rational consumer.
1.1 Research Question

Two important issues have been addressed above. The first issue is that a growing body of knowledge seems to suggest that many decisions are made beneath the level of cognitive consciousness (Bargh, 2002; Camerer et al, 2005). The second issue is that much market research is conducted in ways that assumes that the first issue does not exist. These two aspects, namely the existence of such nonconscious processes and the elusive nature of them which inhibits people from being able to report them, is often neglected in traditional consumer research and theory.

Considering the many conflicting messages consumers are exposed to daily, it is recognized that it is not unlikely ads affect each other. Consumers might perceive advertisement differently when being exposed to them in a context of other advertisements or in a specific sequence. Because many of the processes assumed to influence judgement have proven to occur under the threshold of cognitive consciousness, this sequential or contextual spill over effect might never enter the conscious awareness of the consumer, thus influencing judgements and decisions in ways that likely cannot be reported.

Hence, investigating this sequential and contextual spillover effect would require other tools than the techniques and methods provided by traditional consumer research. Modern neuroscience might supply modern marketing research with methods offering more insight, or the very least, a complementary picture to what the traditional research methods within this field might provide.

The foundation and motivation of this paper thus relies on the two conflicting ideas paired with the availability of new tools and ways of examining advertising effects. Consequently, acknowledging the above mentioned challenges, the main essence of this thesis can be summarized in the following research question:

“How are consumers’ nonconscious reactions and conscious judgement of ads affected by the context and sequence of which the ads are presented in?”

The goal of this paper is thus to investigate if advertisements effectively should be researched in isolation, if consumers really are able to sort out what information is relevant for judgement and what role nonconscious reactions might play
1.2 Reader’s guide

To answer the research question of this paper and to provide for an intuitive graspable framework the following is an illustration and explanation of the structure of the thesis.
Following the above pictured reader’s guide, a brief account of the epistemological standpoint and most relevant delimitations within this thesis will be provided.

Following this, a discussion and literature review will be held with the focus on 3 key theoretical propositions that all intertwine with the research question; A section on traditional marketing theory and research, a section on more modern views on consumer behavior and a section on the contributions of neuroscience. Lastly, a section is devoted to discussing variable definitions.

Following a summary of the theoretical framework, hypothesis will be defined and based on these, a research design framework will be suggested.

After describing the research design framework, the defined hypotheses will be tested and all results will be displayed. This will be followed by a reflection of experimental quality and what specific limitations to consider. The thesis will conclude with a broader discussion relating the findings both to relevant academic theories as well as suggesting the related marketing implications.

Finally, a conceptual framework integrating the findings with existing theories will be suggested.

### 1.3 Epistemology

In order to provide for a sound reasoning regarding the methodological approach for answering the research question of this thesis, a brief discussion regarding the epistemological standpoint will be held.

*Epistemology* can be described as the question of what can be considered as valid *knowledge* within a certain field. The scientific philosophy adopted will guide the choice of methods, the data collection, the interpretation of results as well as subsequent theorizing based on the findings (Bryman & Bell, 2005).
The underlying epistemological assumption guiding this work is that of a positivistic perspective. This is due to a number of different, but many times related, reasons. In essence, a positivistic approach to science means that only phenomenon that can be observed through the senses or are drawn from logical or mathematical reasoning counts as valid (Bryman & Bell, 2005). Traditionally, studies conducted in a positivistic manner are of quantitative character as opposed to more qualitative studies. However, quantitative studies and methods are not exclusive to the positivistic philosophy.

Furthermore, other positivistic notions include the idea of objective (as opposed to normative or interpretive) science paired with the idea that the researcher should (and are able to) be completely unobtrusive and not affect the object of research in any way. The role of theory is to provide reasonable hypotheses that can be tested in a way that allows the researcher to postulate some form of causal relationships (Bryman & Bell, 2005). It is also assumed that the results of such positivistic research are dissociated from person and character (i.e it does not matter who the researcher is, the result should be the same) and that concepts such as reliability and validity are central in the construction of a good research design.

More recent perspectives, dominant within many of the social sciences, such as interpretivism and constructionism are often portrayed as the antithesis of the above presented positivistic perspective (Bryman & Bell, 2005). These perspectives acknowledge different subjective realities and point out that the methodology and tools related to the natural sciences perhaps offer limited insights when conducting research with humans because humans, unlike for example, atoms, are capable of ascribing meaning to their surroundings. They share the concept of science as pointed out by the likes of Max Weber (1947) and Alfred Schutz (1962) in that the social sciences should not strive to define and explain causal relationships as the natural sciences does, but instead strive to describe and understand human behavior.

These perspectives reject many of the underlying assumptions that underpins the positivistic concept. It is assumed that since humans ascribe meaning to their surroundings and that all humans are separate individuals that there is no such thing as an objective research practice. Furthermore, the focal point of interest usually also differs. Acknowledging that humans create and live in
separate subjective realities, phenomena existing within these realities such as interpretations and feelings become valid objects of research.

The nature of this thesis raises some interesting epistemological concerns. On one hand, it strives to understand and describe social phenomena such as advertising and, on a larger scale marketing, while on the other hand investigating biological stimulus-response models utilizing strictly positivistic means and methods. However, the different stances described above represent two ends of a continuous spectrum of different perspectives and related practises. While the underlying assumption and goal is that of an objective research practice, it is acknowledged that such a thing as a completely objective research practice is perhaps only attainable in theory.

However, the experiment is designed in a way that strives to be as objective as possible (see chapter on experimental design). By relying on validated and frequently used methods, the intention and ambition is to make the study as reliable, reproducible and free from subjective bias as possible. In line with the adopted positivistic perspective on research, this study is best described as a quantitative study as the data collected will be analysed using traditional quantitative, statistical methods.

Closely related to this discussion is the distinction between deductive and inductive workflows. Bryman & Bell (2005) states that a deductive workflow starts with theory out of which hypotheses are formulated. This will guide the data collection and analytic efforts. The hypotheses can subsequently either be confirmed or rejected, potentially revising the existing theory. The inductive workflow takes its departure from observations in formulating theories. In other words for inductive work theory/data comes first, while for deductive work theory is the starting point.

This thesis is best described as a deductive work with elements of inductive reasoning. The foundation for the thesis is provided by a wide body of existing theories and these have guided the data collection, a procedure compatible with a deductive workflow description. On the other hand, the result and analysis of the conducted experiments has allowed for new perspectives, guiding subsequent theorizing, vaguely akin to an inductive workflow.
1.4 Delimitations

One of the main focal points of this thesis is based on the notion of unconscious consumer behaviour and emotion. Traditionally, emotion has been described as consisting of 3 components - valence, arousal and motivation. In the limited scope of this paper it has been decided to focus on valence and arousal and thus motivation will only briefly be touched upon.

Furthermore, it was decided to focus solely on print advertising as this provides several benefits over other types of marketing tools such as TV-commercials. In terms of distortion in the produced eye-tracking data, print advertising, should carry significantly less potential sources of distortion than TV-commercials because there is no need for the pupil to refocus on moving objects or react to differences in luminance etc. Furthermore, moving commercials often contains sound which also has shown to have an effect on pupil dilation (Gorn et al. 2001) - another source of potential distortion.

Focusing on print ads also allowed for a relatively short experiment, something that was deemed important to keep mental workload, fatigue and motivation at appropriate levels.

While it is very plausible that any observed effect could potentially apply to other media like auditory communication, TV-ads, etc. it was necessary to limit the scope of this work. Thus, to reach any conclusion on contextual effects affecting other forms of media, those would have to be investigated separately.

There is extensive literature relating different cognitive and emotional processes and their effects within certain product categories (Kover & Abruzzo, 1993; Vaughn, 1980). However, as investigating the highly complex nature of potential contextual effects is an extensive pursuit in its own right, including product categories as another variable to consider would have required the research design to be significantly more complicated. Because of this, it was deemed more meaningful to not include discussions regarding product categories.
1.5 Concept definition

As will be demonstrated, a growing body of knowledge seem to suggest that it is perhaps not always the cognitive, rational and conscious processes that controls our judgements and behaviors. Emotions, moods, feelings and related concepts are gaining increasing attention as crucial drivers for judgement and behavior. However, these concepts traditionally have been described rather abstract and often differently across various academic disciplines. The terms emotion and feelings are, for example, often used synonymously and are to most people just two different words for the same concept. Therefore, in order to provide for graspable and clear variables for the purpose of this thesis, a distinction between emotions and feelings needs to be made.

1.5.1 A brief history of definitions for “Emotions & Feelings”

From the perspective of evolutionary biologists and psychologists (such as Darwin and James), emotions are primarily a biological phenomenon (such as bodily responses) (Cornelius, 1996; Genco et al, 2013; Oatley et al, 2006). More contemporary work have highlighted both cognitive aspects, especially appraisals, of emotions as well as the idea that many emotions are socially constructed and serve social purposes (ibid).

In his book The expression of the emotions in Man and Animals (1872) Darwin posed two broad questions - How are emotions expressed and where do they come from? Emotions are here described as basic survival mechanisms. Darwin described a link between the emotion, the expression of that emotion and the bodily system involved (Oatley et. al, 2006).

William James (1890, p.449) discussed emotion according to the following logic; “... the bodily changes follow directly the perception of the exciting fact, and that our feeling of the same fact changes as they occur is the emotion”

James thus suggested that if there was no bodily changes, there would be nothing left to call “emotion” (Cornelius, 1996) and when we feel frightened, we do so because we become aware of
the fact that our heart is beating harder, our posture freezes or our breathing intensifies (Oatley et. al, 2006).

In the mid-20th century, many social constructivists started to reject the idea of emotions as primarily a biological and physiological matter. Influential authors for this era such as Arnold & Gasson described emotions as “The individual must perceive and judge the object in relation to himself (as suitable or unsuitable, as good or bad in relation to himself) before an emotion can arise” (Arnold & Gasson, 1954, p 294).

Here, emotion is more dependent on some form of conscious judgement, or appraisal (and this line of reasoning has widely been labeled appraisal theory). In contrast to James (1890) perspective, Arnold & Gasson (1954) thus describes an almost reversed emotional sequence where conscious appraisal in part modulates emotion and bodily responses.

Social psychology has also contributed to the understanding of emotions. In most models descendant from this academic discipline, the term affect is used as a label to describe both emotions and moods (Forgas, 1995). Forgas (1995) claims that Emotions are more intense, short-lived and usually have a definitive cause and a clear cognitive content (e.g anger or fear). While this definition share some features with older perspectives on emotions (such as Darwin’s), Forgas (1995) thus also includes a cognitive component into the concept of emotions.

In the mid-20th century Walter Cannon (1927) suggested that if the linear emotional sequence suggested by James holds true, one could effectively reduce the level of emotion by severing the part of the brain called the viscera (which is where James suggested that bodily emotions arises). However, through various examples and studies, he could show that this procedure, in fact, resulted in more intense emotions and not less emotions as James would suggest.

Following Cannon, Hess (1950) suggested that emotions seem to depend on programs of electrical impulses based in the brain (Oatley et. al, 2006).
Figure 1 is a summarizing table over some of the most important authors trying to define the term emotions:

<table>
<thead>
<tr>
<th>Year</th>
<th>Author</th>
<th>Definition of emotion</th>
</tr>
</thead>
<tbody>
<tr>
<td>1872</td>
<td>Charles Darwin</td>
<td>Survival mechanisms, evolved through natural selection, reflex-like automatic response</td>
</tr>
<tr>
<td>1890</td>
<td>William James</td>
<td>“bodily changes following directly the exciting fact... and feeling of the same changes as they occur IS the emotion”</td>
</tr>
<tr>
<td>1943</td>
<td>Walter Cannon &amp; Walter Hess</td>
<td>Electrical signals from neuron to neuron following well-coordinated patterns</td>
</tr>
<tr>
<td>1954</td>
<td>Magda Arnold &amp; J. Gasson</td>
<td>“An emotion can be considered as the felt tendency toward an object judged suitable, or away from an object judged unsuitable”</td>
</tr>
<tr>
<td>1994</td>
<td>Antonio Damasio</td>
<td>“It is an expression of basic mechanisms of life regulation developed in evolution and is indispensable for survival... they are bio-regulatory reactions that aim at promoting survival and what we as conscious and thinking creatures identify with well-being. They are constituted by patterned collections of chemical and neural responses that the brain produces when it detects the presence of an emotionally competent stimulus.”</td>
</tr>
</tbody>
</table>

Figure 1

1.5.2 Modern definitions of “Emotions and Feelings”

Within modern social psychology and neuroscience, the term “emotion” is usually described as encompassing three primary concepts - *arousal* which relates to the intensity or strength of the emotion, *valence* which dictates the direction of the emotion from strongly positive to strongly negative and *motivation* which is the action orientation component of the emotion, specifying the related behavioral response from approach to avoidance (Genco et al, 2013).

Emotions in this light are often described as nonconscious automatic responses. These reactions are manifested in autonomous responses in the nervous system, behavioral responses and glandular responses. Autonomous responses are commonly changes in blood pressure and increased heart rate; behavioral responses are generally freezing, withdrawal or approach; glandular responses are
commonly general sweating and hormonal reactions (Baars and Gage, 2010; Hansen and Christensen, 2007). These emotional responses have been found to be extremely important drivers of all human behavior - consumer behavior included (Genco, 2013).

On the other hand, “feelings” are still most often describes in line with how Damasio (1994) described them; the conscious experience of being in an emotional state (Genco, 2013). Based on this view, feelings are derived from emotions, but with the important difference that they are now subjectively and consciously experienced.

Traditional market research has mainly been concerned with feelings partly because from traditional market perspective there has not really been a need to distinguish between emotions and feelings and partly because there has not been an intuitive and accessible way of measuring emotions.

Seeing that more recent perspectives on consumer behavior acknowledges that emotions are an important part of the process in consumer behavior, it is problematic that the research methods affiliated with the traditional perspectives seemingly cannot measure them adequately.

This thesis will rely on the three-component view of emotions dominant within social psychology and modern neuroscience. Hence, emotions are defined as non-conscious automatic bodily responses while feelings are the subjective experience of such emotions.

Because arousal and valence are by far the two most widely discussed and used concepts, these will serve as master variables for the experiments conducted within the framework of this paper. An in depth definition of these two terms as well as means of measuring them is offered in later chapters.

A final, and important distinction that is often highlighted when discussing the differences between emotions and feelings is the duration. Emotions (as they are defined above - automatic responses) arise and diminishes within a few seconds while feelings on the other hand is much longer lasting (labeled self-reported aspects of emotions in the schema below) (Oatley et al. 2010).
While *arousal* belongs in the very first segment of the emotions-timeline below, *valence* can, to some extent, extend into consciousness and be subjectively reported. For this reason, valence has in several studies been measured through subjective rating scales (such as Likert scales labeled “Pleasant-Unpleasant” etc.) (Gorn et al, 2010; Feldman, 1995; Bradley et al., 1994; Bradley et al., 2008; Aaker, 1986; Lang, 1995).

Figure 2 illustrates the time-spectrum of emotional reactions.

![Figure 2](image_url)
2 Traditional marketing theory and research

This theory chapter will commence with a brief description of the different, traditional perspectives on consumer behavior. These will be accompanied by a description of the related, and for this thesis relevant, marketing and communications models based on assumptions of the traditional perspectives.

2.1 - Traditional perspectives and research of consumer behavior

One broad framework for understanding the evolution of consumer research comes from Østergaard and Jantzen (2000). In essence, it is suggested that there has been 4 shifting and gradually overlapping perspectives during the last 50 years that has had significant effect on the associated practices in marketing. Below is an illustration of the 4 perspectives and their differences. A more detailed description will follow below.

2.1.1 Buyer behaviour

In the early 60’s, buyer behavior was the dominant perspective. Research was highly reliant on laboratory experiment where all variables could be controlled as this very behavioristic concept focuses on the ongoing stimuli-response process humans were perceived to be constantly going through. The metaphorical concept used to describe the consumers is often an animal as consumption was assumed to be an instinct driven response to whatever need emerged in the consumer and how different stimuli could be used to affect these needs. The perhaps most famous example of this in an experimental design is Pavlov's theory on conditioning (Pavlov, 1927).

While this paradigm assumes the consumer is uncritical and ignores other factors of importance (that the following perspectives start to incorporate) it still has its place in theories of purchase
behavior. Below, a table presenting an overview of the relevant consumer research paradigms is offered (adapted from Østergaard and Jantzen (2000)).

<table>
<thead>
<tr>
<th>Perspectives</th>
<th>Buyer behavior</th>
<th>Consumer behavior</th>
<th>Consumer research</th>
<th>Consumption studies</th>
</tr>
</thead>
<tbody>
<tr>
<td>The consumer metaphoretically</td>
<td>Animal</td>
<td>Computer</td>
<td>Tourist</td>
<td>Tribe member</td>
</tr>
<tr>
<td>Ontology of consumption</td>
<td>Mechanical instinct driven</td>
<td>Electronic rational</td>
<td>Emotional narcissistic</td>
<td>Metaphysical symbolic</td>
</tr>
<tr>
<td>Character of the subject matter</td>
<td>The stomach's need</td>
<td>The brain's wants</td>
<td>The heart's desire</td>
<td>The eye's recognition</td>
</tr>
<tr>
<td>Primary method</td>
<td>Experiments</td>
<td>Questionnaire</td>
<td>In-depth interviews</td>
<td>Fieldwork</td>
</tr>
<tr>
<td>Scientific foundation</td>
<td>Behaviorist physiology</td>
<td>Cognitive psychology</td>
<td>Existential psychology</td>
<td>Cultural and social theories</td>
</tr>
</tbody>
</table>

### 2.1.2 Consumer behavior

What changed with the emergence of consumer behavior (as opposed to the previous perspective) was the recognition that something happens in the consumer between the stimuli and the response.

Several studies were made on the process of how consumers receive and interpret information from different environmental stimuli. While buyer behavior research had taken place in a closed controlled environment, the researcher now tried to get closer to the consumers in their own natural environment. The emerging theories were most commonly tested through surveys and questionnaires with the most prominent assumption being that consumers would act rationally and were able (and willing) to tell the “truth”.

The metaphoric shape of the consumer in this perspective is that of a computer. The major criticism of the buyer behavior perspective is the notion that behavior is explained through some kind of genetic coding. In consumer behavior, instead, information is argued to be processed by the consumers in a rational way and from thereon result in the manifestation of attitudes and beliefs
that will dictate behavior. One of the dominant theories created under this perspective is Fishbein's theory of reasoned action (Fishbein, 1980)

2.1.3 Consumer research

With consumer research, a significant change took place in the assumptions about consumers – often described as the interpretive turn (Sherry, 1991). The idea that consumers were, and acted, completely rational was now rejected. Instead the assumption was that ‘irrational’ emotions and feelings were important drivers in the decision-making process.

Where buyer behavior had consumers trying to fulfill some kind of need, and consumer behavior trying to maximize the perceived value of beliefs, attitudes and attributes (Fishbein, 1980), the assumption of irrationality now meant a change in the perspective towards consumption as being a desire for a meaning in life (Østergarrd, 1991).

As it was still an assumption that all consumers must operate in a socially constructed world, the research now shifted towards an attempt to understand the individual's perception of this world and their desires. Research methods therefore now focused on in-depth interviews to gain further insight into the individual although the perspective of consumer research still is similar with consumer behavior in the way that the focus is in the individual – the focal difference is the assumption about rationally - or lack thereof.

2.1.4 Consumption studies

The 4th perspective was consumption studies, which offered a significantly different perspective than the previous three paradigms in how consumers were perceived. Where research had previously focused on the individual, consumptions studies argues that consumers should be studied as tribe members (Maffesoli, 1996) that collectively create meaning based on similar perceptions of the universe they exist in.
The symbolic meaning that consumers attached to different stimuli was now seen as way of trying to gain recognition of the tribe and its members rather than the more narcissistic motivations that consumer research was based on. The main research method in this perspective was fieldwork where research takes place in the tribes setting and the tribe’s creation of meaning/symbols is investigated rather than the individual.

2.1.5 Perspective of this paper

A main theme of this thesis is to highlight the inadequacies of traditional marketing research and thus it should not come as a major surprise that the perspective of the thesis does not really fall into any of the four traditional perspectives on consumer behavior described above.

While it can be appreciated that perspectives in an increasing degree has acknowledged that emotions and feelings plays a role in consumer behavior, the lack of understanding of unconscious behavior is still significant. Furthermore, in terms of the definition of emotions (as defined in this paper), it is a paradox that the later perspectives of consumer behavior acknowledge that emotions play a role in the consumer's behavior while still mainly using self-reporting methods to investigate these emotions.

Accordingly, this thesis would be placed in what could be called a new paradigm. In this perspective, consumers are still viewed as emotional creatures that are inherently irrational. The ontology of consumption would be in line with ideas of consumer behavior studies, but in order to investigate consumers, however the primary methods would be through experiments and measurement of physiological variables which can be dated all the way back to buyer behavior that in contrast assumes complete rationality.

Having established the previously dominant perspectives on consumer behavior, an overview of selected marketing and communications models and theories, related to those perspectives, will be presented below. Following these, a contrasting picture suggested by more recent insights regarding consumer behavior is presented.
2.2 Traditional Advertising & Marketing communication models

Traditionally, the function of advertising has been described through various hierarchy-of-effects models originating from the field of marketing and communication such as the AIDA (attention-interest-desire-action) and the ELM (elaboration likelihood model) (De Pelsmacker et. al, 2010). While the AIDA model assume a linear chain of effects and claim a cognitive-affective-conative sequence, the later proposed ELM model acknowledges that perhaps certain messages are processed outside of our central cognition (ibid).

Understanding and predicting how consumers react and behave according to advertisement is a crucial ability for marketing managers. Numerous different models and theories have been designed to handle this task have been developed over many decades. While one conclusion is that a single model to explain everything might not exist – but rather that several different models could explain different scenarios – some models have been more accepted and applied than others.

Below is a description of the more notable and widely used models and how they might fail or help in answering the research question of this paper.

2.2.1 Hierarchy-of-effects models

Hierarchy-of-effects models date all the way back to the 1890’s, making them some of the oldest and most influential models on marketing thinking. Around the 1980’s the models started losing support due to certain assumptions that have come into question, some of which has been discussed in the evolution of consumer behavior theories. However, a quick look into the mechanisms of such models is important, as it has laid the foundation for other models that have followed.

Hierarchy-of-effects, essentially, as the name might reveal, assume that a process must happen in a specific order and a latter step cannot happen until the previously referred steps have. Common to all models within this framework is that consumers will go through 3 different stages in
responding to advertisement or other marketing communication (Palda, 1966). The common three stages that are present in most forms of such models (in one form or another) are:

1. The cognitive state, in which consumers engage in thinking, leading to awareness or knowledge of the brand/product/idea/etc. being communicated.

2. Affective stage, in which emotions and feelings associated with the brand/product/idea/etc. is formed.

3. Behavioral (or conative) stage, in which consumers take action in regards to the communication and change behavior – for an example, by buying the advertised product.

The Hierarchy-of-effects models all assume that consumers go through this linear line of steps and the majority of models claim these steps are the above mentioned cognitive-affective-behavioral sequence (DePelsmacker, 2010).

2.2.2 AIDA & FCB

There are many models fitting the description of an Hierarchy-of-effects model, but to mention a few, the most frequently referred to in marketing literature might be the AIDA (or revised versions of this) and The Lavidge and Steiner model (DePelsmacker, 2010).

The AIDA suggest 4 steps - Attention, Interest, Desire and Action while The Lavidge and Steiner model suggest the steps of Awareness, knowledge, liking, preference, conviction and purchase (ibid). The main difference between these 2 models and every other Hierarchy-of-effects model is the disagreement about the sequence of the different stages – and the exact labeling of the different stages.

Expanding on the reasoning of the hierarchy-of-effects models, Vaughn (1980) suggests the Foot-Cone-Belding grid (FCB), incorporating a situational aspect. He suggests that 4 different situations can be distinguished on 2 dimensions – the low-high involvement, and the think-feel dimension.
The FCB grid was suggested as a more context sensitive alternative to the rigid Hierarchy-of-effects models.

The high-low involvement dimension is defined as the importance people attach to a specific product or the buying decision. This could involve financial risk or social risk just to mention a few. A candy bar would likely not entail a big financial nor a social risk (unless the consumer is heading to a weight watchers meeting perhaps) and will under the FCB framework be seen as a low involvement type of good. However, a mortgage loan or a new car would likely be connected with a substantial amount of risk and would instead be a high involvement type of decision.

The think-feel dimension is a reflection of to what extent the decision is based on cognitive or affective motives. To illustrate - buying milk might not involve very much affective motives while perfumes might to a higher degree do so, and downplay cognitive motives.

Having defined the four possible quadrants of the model, Vaughn (1980) then put up different suggested orders of the cognitive-affective-behavioral order – or as he calls them Think-feel-do, illustrated in Figure 3.

![The FCB-grid](image-url)

While the FCB grid were a bit more sensitive and took more factors into account than the previous hierarchy-of-effect models, all of these models are subject of some fair criticism and are today being less used practically in marketing communication. One important criticism is that no
empirical research has been able to find any specific hierarchy of cognitive, affective and behavioral steps that consumers are assumed to go through (DePelsmacker et al. 2010).

Another criticism is the lack of interaction between the different stages of the model. In real life this is likely not a realistic carved-in-stone sequence, as purchase will affect experience, which will then affect attitude etc. The complexity of the consumer decision process is simplified to an extent that has severe problems when faced with describing and explaining the real nature of consumption.

Finally, in relation to this paper, another problem that the models do not address is the environment of the advertisement. The models exclusively analyses a single ad (thus, in isolation) and neglects to address the facts that these steps might be significantly affected by other factors externally.

Commenting on the limited practical use of hierarchy-of-effects models in modern marketing, Barry & Howard suggests in their review of the hierarchy-of-effects models that “...viewing the hierarchy of effects simply as a heuristic model which may have utility for general planning and guidance purposes may be most appropriate” (Barry & Howard, 1990).

Despite these limitations, the models are important ones as they through their framework has led attention within the field towards attitude formation and help shape some of the models that are widely used today.

2.2.3 Elaboration likelihood model

Around the 1980ies, focus started to shift towards attitude as a crucial parameter in marketing communication. Attitude can be defined as “person’s overall evaluation of an object, a product, a person, an organization, an ad, etc.” (DePelsmacker et al. 2010).

As discussed in the Hierarchy-of-effects models, a limitation is that people react differently and no single model can explain everything. Trying to solve this problem, the Elaboration likelihood model (ELM) appeared.
ELM looks at the way attitudes are formed on a cognitive, affective or behavioral level, while the second parameter is about the level of elaboration of a message happening - or central-route versus peripheral-route processing (Petty & Cacioppo, 1986). An illustration of the ELM schema can be found in appendix 1.

Under the central processing route, attitude formation will be a result of careful and thoughtful process where the consumer has evaluated the pros presented in the ad. This route involves a high level of cognition and it is required by the consumer to understand the argument posted in in the ad. It is predicted that the resulting attitude from this route is relatively enduring and long lasting.

However, another route, the peripheral route, is also suggested where attitude formation is based on a consumer's association with different cues in the stimulus. These cues do not have to be logically connected to the quality of the stimulus.

The ELM is very similar to the FCB, but some major differences is present. First of all, no linear sequence of events is assumed, which was a clear limitation of the Hierarchy-of-effects models.

Secondly, the ELM model acknowledges that perhaps certain messages are processed outside of our central cognition, which is relative drastic change from traditional marketing theories (Pelsmacker et al, 2012) – but incredibly relevant to this paper and its research method.

However, the ELM also has some limitations. First of all, while the model does incorporate emotion and feelings in its framework, its distinction between the two concepts is not great. Furthermore, while the model argues that peripheral cues can alter attitude to some extent, it, like many other traditional models, does not consider the impact of a context. Effectively, ads are analyzed in isolation.
3 Modern views on consumer behavior

To contrast the above reviewed models and perspectives, below is a brief account of more recent perspectives on consumer behavior. This review will include investigating the concepts of priming, arousal, valence and the relationship between the two latter concepts. Furthermore, important theories and models primarily from social psychology will be incorporated into a theoretical framework leading into more practically oriented methods chapter, designed to answer the research question.

As has been argued, many of the older perspectives on consumers and their behavior has proven to be inadequate in fully accounting for the complexity related to the subject of this thesis. What most of the previous models of consumer behavior have in common is that they assume that the consumers in one way or another act rationally (while some models assume economic rationality where consumers always seek to maximize their economic utility, others might assume some form of social rationality where consumers rationally seek the social approval of others). It has usually been assumed that consumers behaved rationally in relation to the information they had at hand and what they wanted to achieve, that their choices were consistent with their preferences and that these preferences were unambiguous and stable unless they were presented with a better, more rational alternative (Genco et al, 2013).

However, anyone who has ever interacted with another human being would likely be hard to convince that this is an accurate description of human behavior. Consumers repeatedly behave in ways that violate most of the expected outcomes of the many different rationality based consumer behavior models that has had supremacy in the academic field of consumer behavior for the last century. Rarely can consumers fully account for the rationale behind their decisions or preferences and seldom can they reliably predict their future actions and choices.

As suggested previously, a growing body of knowledge seems to suggest that emotions play a much greater role in consumer behavior than it has been previously assigned. Lee et al. (2009) even claims that emotions are the best predictor for consistent preference and behavior.
3.1 Contributions from social psychology

How emotions affect cognition and judgements has been a prime point of research within social psychology for many years. While little progress has been made in terms of creating a unifying cross-disciplinary framework, most researchers agree on two things - 1) emotions affect cognition and 2) cognition affects emotion (Kuppens et al. 2012).

As the goal of this paper is to investigate the effect of context and sequence on emotion (by looking at arousal through measuring the dilation of the pupil) and on evaluative judgement (by having subjects reporting their cognitive assessments of the ads) models dealing with these concepts are certainly relevant to discuss.

As described in the discussion regarding the distinction between emotions and feelings, there has been a general inconsistency across different academic disciplines in how feelings and emotions have been used and described. Furthermore, the relation between the two concepts and in turn, how they relate to the concept of affect is also somewhat unclear. According to some neuroscientists, the term emotions is described as synonymous with affect or affective states (Genco et al. 2013). Other authors such as Pham et al. (2010) discuss affective feelings as encompassing “moods, emotions and affective reactions”.

While this paper is devoted to the more neuro-oriented definition of emotions, a number of the prevailing models from fields such as social psychology incorporates many of the components of how emotions have been described in this paper (such as automatic bodily responses) into their definition of feelings. This leads us to conclude that theory or models dealing with feelings should necessarily not be discarded as being outside of the interest of this paper solely based on this linguistic difference.

With this in mind, a closer inspection of the various theories and models describing the role of emotion and feelings in judgement is motivated. Very closely related to the research question of this paper is the view that “… when in a positive or negative emotional state, feelings are likely to color your evaluative judgements of events and objects as good or bad, even when the objects beings judged have no relation to the cause of the emotion” (Oatley et al. 2010, p.279).
One academic discipline (from which the above stated quote is retrieved) that has had profound impact on the evolution of perspectives on consumer behavior and later also more specialized fields such as neuroscience and neuromarketing, is social psychology. Social psychology has provided a number of models for explaining human (and thus consumer) behavior in various circumstances that is still being widely taught and used. Insights from this field has been applied extensively within neuromarketing and is perhaps most relevant for understanding of how conscious and non-conscious processes interplay and affect human (and consumer) choice and behavior (Genco, 2013).

Recent developments within social psychology have also suggested that affect and emotions might have a larger impact on the human decision-making process than previously assumed (Kahneman, 2011, Pham, 1998,). This insight, amongst a number of other factors, has been one out of many arguments as to why, for example, older methods of conducting consumer research might prove to be inadequate.

### 3.2 Priming

The focal point of interest of this thesis is how consumers might be non-consciously affected by the sequence and context in which they view ads. The notion that 2 completely unrelated ads could affect each other might intuitively seem unlikely, but the inspiration for uncovering this question is based on a very similar concept that has been well proven to exist and affect consumers on a non-conscious level - the concept of priming.

In traditional marketing view, the goal is often to persuade consumers to perform a certain behavior – i.e. buying the product. Persuasion in this regards, has traditionally been claimed to require a series of steps including getting the consumers attention, getting them to judge it, find it reasonable and remember it - a rather linear chain of events, not too different from the described Hierarchy-of-effects models. As these descriptions of the consumer journey has come under criticism, priming has gained a considerable amount of attention due to the fact that it requires none of the above-mentioned steps to affect behavior.
Priming is based on the mental processes of associative activation, which is the brain's ability to automatically, non-consciously and quickly find associated ideas that relate to whatever idea the consumers is being exposed to (Genco et al. 2013). For example, if you are asked to think about the idea of a birthday your brain will immediately look into its massive networks of similar ideas that relate to birthday. These associations never reach your consciousness but they are stored ready in your mind. This reaction is a mechanism that enables the brain to prepare you for any given situation.

What is referred to as emotional markers or somatic markers (Damasio, 1994) is the brain's ability to retract memories of bodily experiences from the past. Regardless of the intensity and valence (negative or positive), these experiences are encoded into the brain and without consumers ever being aware, and these can be triggered by primes. As soon as this ability was discovered it didn’t take long for marketers to realize the potential it had in commercial context.

The role of emotional markers is big in consumer judgment and choice. Even though we might not be aware of it, consumers mark different brand/products with labels of attraction/aversion and these labels supply a quick shortcut to evaluate those product/brands we faced with a decision.

According to Damasio (1994), these markers are updated continuously through every experience in the consumer life. This is also a source of inspiration to the framework of value-based choice suggested by (Plassman et al. 2012).

Priming is the associative activation. When one idea has activated another idea, it would be said that the first ideas has primed the second. If priming is done well it can even trigger behavior (consumption/purchase) in which case we would say the purchase was primed by the initial stimuli (Genco et al. 2013).

Priming can be achieved through any number of setups. Below are a few examples:
Using subliminal priming, the activation of certain goals were achieved by Chartrand (Chartrand et al. 2008). Two groups were exposed to words associated with either “prestige” or “Thrift” during a short task. Following the task, participants were asked to make a hypothetical choice and it was found that subjects were more likely to choose socks from Nike at $5.25 a pair when exposed to the “prestige” condition than Hanes $6 for two pairs when exposed to the “thrift” condition.

Another famous experiment showed how priming people with certain categories of words could alter their behavior unconsciously. In an experiment by Bargh (1996), two groups of young students were asked to create 4 word sentences from a selection of 5 words (for example; how, finds, it, yellow, instantly). One group was primed by being exposed to words associated with elder hood (for example, wrinkle, bald, grey). When the participants were asked to walk to another room for another experiment following this exercise, the subjects who had been exposed to the “old” prime walked significantly slower to the next room.

Quite interestingly, the reverse effect also exists. In a German study, researchers tried to reverse the effect of Bargh’s study. They asked participants to walk around a room for minutes at a rate of 30 steps per minute. That translates to roughly one third of normal speed. Afterwards participants were much quicker to recognize words relating to elderly, like old, forgetful and lonely than had they not been asked to walk slower than normal (Kahnemann, 2011).

Advisement, product placement, store display, etc. could all be viewed as primes that are aimed to influence consumers. The discipline of branding can essentially be described as one big exercise in priming as the goal is to create associations between the product/brand and any pleasant/unpleasant (dependent on the product) ideas or feelings.

Priming is usually categorized into 2 different types;

### 3.2.1 Associative priming

This type is short lived and is often relative weak. Associative priming is not related to any internal goals of the primed subject and therefore the prime can easily decrease over time or if it is
interrupted. A good example is the previously mentioned study where participant’s walking speed was dependent on what category of words they had previously been exposed to (Kahneman, 2011).

Illustrated here is a possible associative priming example. If one group of subjects were exposed to the left column of words, priming theory would suggest that chances are very high that they would fill in the empty line with the word “green” or “grey” because they have been primed with the concept of colors. Another group of subjects exposed to the right column of words would instead be likely to fill the empty line with the word “grape” as they have been primed with the concept of fruit.

### 3.2.2 Motivational priming

This type relates to subject’s goal and is therefore stronger and more influential on behavior. An example could be a person with a goal to lose weight. If that person is exposed to advisement of slim ladies that drink the same weight-loss drink, this prime could in time trigger an association between that product and the goal of losing weight. Motivational priming often tend to grow stronger over time (branding) and only diminishes when the goal is achieved (Genco et al. 2013).

While priming has been proven to have very strong influences on consumer behavior, it is worth noting that some conditions have to be in place for it to work (Genco et al. 2013). First of all, there needs to be a gap between the consumer’s current state and the goal that is being primed. If the consumer already has achieved the goal that is being primed there will be no effect. Secondly, the goal that is being primed cannot be in conflict with consumer’s existing goals. Trying to affect consumers into eating more meat by using some specific prime will likely not work if the consumer exposed to the prime is vegetarian.

Priming is closely related to the purpose of this thesis. Priming in the context of marketing communication is usually viewed as a deliberate action from marketers in order to persuade consumers into buying their products (motivational priming). What this paper investigates is
whether there is any spillover between these marketing actions (in our framework, those actions are visual advertising) and that these might affect each other (associative priming).

Employing priming-terminology, what is being investigated in this thesis is associative priming from advisement A and how that might affect the effectiveness and impact of advisement B.

Having discussed how unrelated stimuli might affect each other, several models from social psychology can also be applied to help understand this phenomenon. Below is a description of 3 relevant models.

### 3.3 Social psychology models

Three conceptual models, all related to the research topic of this thesis, for understanding the role of affect and emotions will be presented below, Emotional congruence theory, Affect Infusion Model and Affect-as-information (or Feelings-as-information) theory.

While most of these models rely on definitions of emotions and affect that are more diffuse and abstract than those offered by later neuroscientific authors, they are, as shall be demonstrated, still highly relevant for the purpose of understanding the impact and role of emotions on judgements and other cognitive processes.

#### 3.3.1 Emotional congruence

One way of conceptualizing the impact of emotions/affect on cognition and judgement is proposed by Bower (1981). According to Bower’s proposed theory, moods and emotions are comprised of an associative network in our minds (much like the priming effect described). This network connects previous experiences, images, labels and interpretations of sensations in our memory to specific emotions via dedicated pathways. When an emotion is experienced, nodes connected to this associative network will be more readily activated. Bower (1981) then suggests that emotions leads us to interpret, encode and remember our present environment in congruence with our current
emotional state because the nodes in the associated network related to that emotion are more easily activated.

Since our mental capabilities act as a severe bottleneck in receiving, analyzing and storing information, there is a constant selection and prioritization of what information is the most crucial at any given moment (Forgas, 1995). Emotional congruence theory (or Affect-Priming Mechanisms) would suggest that the process of selecting what of the information, presented by the environment, should be attended to is heavily influenced by the current affective state. Thus, Bower states that “affect can prime the encoding, retrieval and selective use of information...” (Forgas, 1995, p.7).

Through a series of experiments, Bowers showed that people were able to better learn and remember tasks and details that were congruent with their current emotional state (Oatley et al. 2010). When people were in a happy emotional state, they were able to better remember positive details in a told story while people in a sad emotional state were better able to remember negative details in the same story.

Other authors have found more inconclusive results when investigating emotional congruence effects. Parrot & Spackman (2000) for example, instead found that memories that were incongruent with the current emotional state were better recalled than those that were congruent. Recent research seems to conclude that mood or emotional-dependent effects do occur, but it is not a mechanism that affects all cognitive and memory related processes as Bower (1981) would suggest.

For the purpose of this paper, emotional congruence is still an interesting concept to discuss. If emotions indeed nudge our interpretation and attention, encoding, memory and judgement in the direction that is congruent with the current emotional state, then there should be some spillover effects of ads in a sequence.

If being emotionally affected by an ad puts the consumer in a certain emotional state, the emotion congruence theory would suggest that the next ad should be interpreted and judged more akin to that emotional state. For example, if a sad ad precedes a happy ad, the happy ad, should, all other things equal, be judged sadder than if it had been preceded by another (or none), happy ad. If this
is the case, it strongly suggests that context and sequence indeed do matter for the way ads are understood, judged and remembered.

### 3.3.2 Affect Infusion model

Another relevant theory derived from social psychology and one that could be perceived to be an extension of Bower’s (1981) theory of emotional congruence is proposed by Forgas (1995) – the *Affect Infusion Model, AIM*.

The model deals with how affective states (emotions as well as feelings and moods) are infused into cognitive tasks and judgements under different circumstances and suggests a number of different possible strategies for forming judgments. It is suggested that; “The AIM assumes that affective states, although distinct from cognitive processes, do interact with, and inform cognition and judgements by influencing the availability of cognitive constructs used in the constructive processing of information” (Forgas, 1995, p.4).

Thus, according to Forgas, if an ad succeeds in having an impact on the emotional state of the consumer (as one could expect that very sad or happy ads might have), it can be assumed that this emotional state has an impact on related cognitive functions such as evaluation and judgement of that ad. However, it is acknowledged that affect might be infused into the cognitive process differently for different kinds of stimuli and situations. Affect might have relatively more or less impact depending on various external circumstances and personal characteristics. This relationship is demonstrated in Figure 4.
The affect-infusion-model

Judgmental target

**Figure 4**

The AIM model outlines four different possible strategies for affect guided judgement formation; *the direct access strategy* which is based on direct retrieval of pre-existing formed judgements, *the motivated process strategy* which is employed when the assessment and judgement is guided by a specific motivation, *the heuristic judgement strategy* meaning that consumers might seek to form judgements through various mental shortcuts and lastly, *the substantive processing strategy* which occurs when people need to engage in selective, constructive processing of information and utilize cognitive processes such as learning and memory (Forgas, 1995). All of these four different strategies for judgement formation comes into play at different situations.

The *direct access* is employed when a previous existing judgment of the same character is readily available. In this case, Forgas (1995) argue that affect should have a small impact and that the previously crystallized judgements is likely robust and resistant to affective distortions. This strategy is by some authors such Oatley et al. (2010) describes as stereotype-like judgements because they are very quick, often pre-determined and bypass most cognitive deliberation of the subject.
By selecting ads that subjects had not been exposed to previously, the ambition was to eliminate such potential bias. This way of selecting stimuli should thus make certain that the direct access route for making judgements and evaluations of the ads were not employed.

The motivated processing strategy is also argued to be resistant to affective influence. This strategy is more goal-directed and is employed when strong motivational and situational desires for a certain outcome exists (Forgas, 1995). Different affect-related goals have been found to drive this strategy – mood repair, mood maintenance, achievement motivation, ego enhancement amongst others. Under these circumstances, the AIM model predicts that affect, mood and emotions has a relatively small impact.

Again, relating this strategy to the research of this thesis, one could assume that under circumstance where subjects desires a certain outcome, effect would likely display a much smaller impact. No such circumstances were expected to be present within the frame of the experiment in this thesis.

A heuristic strategy is suggested for situations where there is no previous experience nor a strong motivational reason to cognitively deliberate on the subject (Forgas, 1995). As humans in general are cognitive misers (Genco, 2013) and will strive to preserve cognitive resources, shortcuts, or heuristics are often used. The AIM suggests that when the object is simple or typical, the personal relevance is low, when there is no specific motivational objectives and when cognitive capacity is limited, then this is the most common strategy. Thus, “judgements may be based on irrelevant environmental variables” (Forgas, 1995).

This strategy seem to be highly relevant for the research question stated in this thesis. The object (ads) is relatively simple, the personal relevance should be low (as measures were taken to choose ads that were unknown to the experiment participants) cognitive capacity might have been relatively limited (as the ads were only shown for a limited amount of time). The AIM would hence suggest that for these circumstances, we could expect external variables (such as context and sequence) to have an impact on judgement.

The last strategy proposed by Forgas (1995) is the substantive processing strategy. This is the most cognitively demanding strategy and arises when subjects are faced with novel or atypical information that is personally relevant and when there is cognitive resources available for
analyzing and interpretation. The AIM suggests that this strategy is primarily used when simpler and more efficient strategies fail or prove inadequate. It is suggested that out of the four described potential routes for reaching judgement the *substantive processing strategy* is the strategy that is the most susceptible to affective impact.

Thus, a counter intuitive suggestion within the AIM that has been proven in several experiments is that the more cognitively demanding the employed strategy is, the bigger the effect from affect, moods and emotions is (Forgas, 1995).

### 3.3.3 Affect-as-information

One of the most influential authors in the field of social psychology, Norbert Schwartz (2011), puts all subjective experiences under the same umbrella when describing the Feelings-as-Information theory; “*Feelings-as-information theory conceptualizes the role of subjective experiences – including moods, emotions, metacognitive experiences and bodily sensations – in judgement. It assumes that people attend to their feelings as a source of information*” (Schwartz, 2011, p.2). One could thus draw the conclusions that according to the feelings-as-information theory, emotions would be one out of a number of factors that has an impact on consumer judgements.

Pham (2004, p.2) states that; “... *they [Schwarz & Clore] suggest that people use their momentary feelings as actual sources of information. People interpret positive feelings as evidence of liking, satisfaction, well-being and so on, and unpleasant feelings as evidence of disliking, dissatisfaction, misery and so on. In doing so, they may fail to recognize that sometimes the actual source of the feelings is not the target being evaluated, but some incidental factors (e.g. good weather)*”

A central theme of the feeling-as-information theory is the HDIF heuristic (how do I feel about it?). According to the proposed theory, people are assumed to ask themselves the HDIF-question when faced with a task involving judgement of some sort. The answer to the HDIF-question is incorporated into the judgement of the stimuli.
However, the idea that feelings and emotions are only incorporated as a basis for judgement and behavior through a deliberate and conscious process is challenged by authors such as Forgas (2001). While these points have been answered by Schwartz (2010) with the valid claim that “people usually consider their thoughts and feelings to ‘about’ whatever is in the focus of their attention, rendering reliance on them the automatic default option”, the affect-as-information theory describes a more conscious and under certain circumstances, more cognitively demanding process of affective impact on evaluations as compared to other models such as the AIM.

Another decisive difference between the affect-as-information theory and the AIM model are the circumstances for when affect and emotions is believed to have a significant impact on judgements. The AIM suggests that the more cognitively demanding an evaluation task is, the more one has to deliberate and analyze the task and the more novel and atypical the available information is, the larger the impact from affect and emotion is. The affect-as-information theory however, suggests that the more conscious deliberation one devotes to the judgement task, the more likely one is to become aware of the true source of one’s affective state and thus be able to disregard those feelings that are not relevant for the judgement.

Having summarized some of the more contemporary theories from social psychology, it has been illustrated how the field increasingly is considering how unrelated and even nonconscious effects might influence consumer behaviour. Since the discussed emotional effects cannot be measured through traditional marketing research methods, contributions from the the field of neuroscience has increased. In relation to this thesis, neuroscience then has some significant insight to offer in terms of answering the research question.
4 The role of Neuroscience

As demonstrated above, consumers are often affected in ways that are far below the level of consciousness. The source of judgement and behavior are seldom clear and accessible to the conscious mind.

There is a vast range of other problems with simply asking consumers about their behaviors as well. Social desirability bias is one, fragmented memory is another, agreeableness bias is a third and the list goes on (Genco et al. 2013). These biases poses serious problems for the way market research usually is conducted since the vast majority of such research revolves around different ways of asking consumers what they think, like, feel, remember, buy and prefer. Genco et. al (2013, p. 11) sums up the discussion by stating that: “The most important of these implications is that human beings actually have very little awareness of why they do the things they do [...] People aren’t lying or trying to deceive the researchers when they make these guesses; they’re literally unaware of the real causes and reasons for their actions.”

If consumers themselves cannot, with an even moderately good accuracy, describe why and what they will be placing in the shopping basket the next time they go to the supermarket, then perhaps it is a fruitless pursuit for the market researcher to ask that particular question. Needless to say, there seems to be a need for other, complementing methods for understanding consumers and their behaviors.

Recently, there has been some significant contributions into the mystery surrounding consumers and their often emotionally controlled consumption behaviors. These insights have, as suggested above, primarily emerged from the academic field of social psychology. Damasio’s (1994) theories of somatic markers, Nobel prize winner Daniel Kahnemann’s (2002) description of the fast and slow system 1 and 2, Schultz’s (1997) insights regarding the underlying neural mechanisms for reward and risk calculations, insights into how emotions affect judgement and behavior (Forgas, 1995; Clore & Schwartz 1983) amongst many others, have all served both to show the relevancy of applying neuroscientific methods and tools within marketing as well as popularizing the topic both within academia and business practice.
Neuroscience can be described as the study of the nervous system that seeks to understand the biological basis for behavior (Plassman et. al, 2012). The last few decades has provided the technological development of certain tools better suited and adapted to specifically study these processes more in depth to provide a more detailed and realistic, or at the very least, complementing picture. The application of these neuroscientific methods and tools (such as functional magnetic resonance imaging (fMRI), eye-tracking, electroencephalography (EEG), facial expression coding and skin-conductance amongst others) in relation to branding, consumer psychology and behavior has emerged as what is being labeled the field of consumer neuroscience and neuromarketing.

This relatively novel field of research seems to provide one promising way of gaining a better understanding of the elusive processes directing consumer behavior. It’s popularity can be illustrated by the explosion of published neuroscientific articles during the first decade of the 21st century (from being a virtually non-existent phenomenon to having over 2000 articles within the field published annually) (Plassmann et al. 2012).

There is an important distinction between consumer neuroscience which is located at the intersection of neuroscience and consumer psychology, and what is being labeled neuromarketing which is the application of neuroscientific tools and insights into marketing, market research and related disciplines (ibid). This thesis will be located middle of that spectrum, using neuroscientific methods and tools to investigate certain aspects of consumer psychology through using real life visual commercial advertisements as the prime stimuli.

4.1 Neuromarketing perspectives on consumer decision-making

As highlighted, traditional marketing thinking often views the decision making process of consumers as a very rational, conscious and cognitive process with little or no impact coming from emotion and non-conscious dimensions.
This view has increasingly been challenged in recent decades and within the field of neuromarketing many studies have increased the emphasis on the interplay between emotions and cognition.

In order to understand the neuropsychological mechanisms of the effects of marketing communication and brands on preferences, the value-based model of choice is an often referred framework (Plassmann et. al. 2012). This model suggests that consumer decision process should be divided into 4 different parts that over a period of time result in preference creation.

The model, illustrated in figure 5, essentially describes the decision-making as a continuous process where the brain is constantly receiving signals and interpreting/evaluating them at each step to create preference. Because it is based on a decade of research in applying neuroscience to the consumer psychology of brands, it considers non-conscious and irrational aspects of consumer decision making to a much higher degree than most traditional marketing communication models.
While this paper is not focused on branding effects specifically, the model is still very much relevant as it does consider what processes and steps are crucial for preference creation. The practical usability of this model, in the context of this paper, is to analyze how the different steps (Representation and attention, Predicted value, experienced value and remembered value) are affected by a context or specific sequence.

### 4.1.1 Representation & Attention

Selecting the information that receives preferential status above alternative information is the mechanism of attention. While consumers are bombarded with stimuli every second, the human mind is not able to process such vast amounts of sensory information, thus letting most of this information go unnoticed (Wilson, 2002).

Because only a fragment of the stimuli is noticed, the understanding of information and how consumers attend it to is crucial in the study of consumer behavior. The first dimension of the value-based model of choice deals with *representation and attention* - which in the context of this paper is highly relevant to the discussion about sequential and contextual effects.

Representation is the first process in identifying a brand. In an example of buying fast food, this involves the processing of information and identifying alternatives – as an example choosing between McDonald's and Burger King. At the same time, the consumer also has to take into consideration internal and external states. An internal state could be the level of hunger while an external could be what alternative a friend chose.

As the main sense of interaction with the outside world, the visual sense also becomes the most important in terms of the incoming information that is constantly being processed. The visual system is highly advanced and allows consumers to identify and evaluate products and brand in matters of seconds (in some cases, milliseconds). In a study, it was discovered that consumers could identify and express preference for 2 different food brands in less than 400 milliseconds, a much shorter time span than would be expected if the consumer would act consciously and do a value-expectancy calculation based on the available alternatives (Milosavljevic et. al., 2011).
What is notable here is that these visual processes in the representation stage are not required to be conscious (Kouider & Dehaene, 2007). This is a key difference from some of the older, reviewed marketing communication theories, and it is an essential difference for the research of this thesis.

According to Knudsen (2007), attention can be divided into 4 components: bottom-up or saliency filters, top-down control, competitive visual selection, and working memory.

*Bottom-up or saliency filters* is automatic and non-conscious and selects the information that is most important from all available information. In terms of visual representation the most commonly named features that has been proven to affect initial eye movement and direct attention is low-level ones like colors, luminance, orientation, size, shape, movement, etc. (Itti et. al., 1998; Wolfe et. al., 2004). All of these factors are combined in a pre-attention scan in order to create an attention map that then designates which information is most likely to be granted most attention – and thus further processed.

*Top-down control* is when internal and external state/goals/expectations influence what is granted attention. The information that is most relevant to a certain goal will acquire the most attention. For an example consumers will pay more attention to beverages if they are thirsty, than other products (Aarts et al., 2001).

*Visual selection* is the identification of all visual areas and choosing which ones are deemed most important (based on bottom-up and top-down attention). Finally, *working memory* impacts how different types of information like vision, audition or somatic sensation compete for entry into working memory networks.

Representation and attention are very complex processes that influence all subsequent steps in the value-based model of choice, and in this thesis the main point of interest is in how representation and attention might be affected by completely unrelated stimulus to the products/brand of interest.

The model suggested by Plassmann et al. (2012) is mostly focused on internal and external states that *relate* to the product/brand when choosing what influences representation and attention. In this paper it is investigated how completely *unrelated* stimuli might also influence representation
and attention – which would be how preceding ads (that has nothing to do with the product, brand or category of the next ad) affects the following ads.

### 4.1.2 Predicted value

The predicted value is the consumer’s expectation of experienced value towards the brand or products. Or in other words, the consumer’s evaluation of how much he/she will have from consuming etc. McDonalds.

Again this process does not need to be conscious but might as well happen automatically beyond cognition. While more traditional marketing models like the Fishbein model (Fishbein et al., 1975) focus on very rational and conscious dissection of beliefs of attribute for a product/brand several studies have shown that “behavioral measures of consumers' positive and negative predicted values for a range of different branded products can be correlated with several different nonconscious physiological responses measured by fMRI amongst others” (Plassmann et al. 2012, p. 7).

In the scope of this thesis, the subjective ratings of valence could be used a measurement of predicted value of that ad/product. For example, if positive ad (for example an ad for ice cream) is rated less positive when viewed in a context, it could be argued that the predicted value of that product has fallen. Although any contextual effects can be related to the predicted value, this aspect is not the focus of this thesis and further research would be need to conclude anything specifically towards this aspect of the value-based model of choice. The main emphasis of contextual effects is related to representation and attention within this thesis.

### 4.1.3 Experienced value, remembered value and learning

The experienced value is the consumer's pleasure from consuming the product/brand and is typically defined as consisting of valence and intensity, while remembered value is the recollection of this value.
Experienced value can be affected by predicted value as many studies have shown that consumers’ belief about objects affect their experienced value (Plassmann et al., 2008). Remembered value can also be altered by earlier steps as one study showed how consumers’ beliefs were altered by advertising when comparing how they recalled their earlier reporting of these beliefs following exposure to advertising (Braun-Latour et al., 2006).

Obviously, the spillover effect being investigated in this paper could also have effects on experienced value, remembered value etc. However, due to the limited scope of this paper, it is the stage of representation/attention that will receive the most elaborate treatment. However, a discussion about how other steps might be influenced by unrelated stimuli will be made.

### 4.2 The role of arousal

As discussed in the review of the evolution of consumer research, the acceptance of emotions influence on decision-making has been increasing since the interpretive turn. It is now mostly accepted that emotional reactions are important driving factors for human decision and behavior, (Damasio, 1994; Lang, 1995; Bradley et al., 1993; LeDoux, 1998). It has been demonstrated that brain damage to areas of the brain that processes emotion can have serious implications on the ability to make decisions and planned behavior (Damasio, 1994).

Arousal also has a significant relevance to marketing and consumer research as it has been argued to be “...the basis of emotions, motivation, informational processing and behavioral reactions” (Groeppel-Klein, 2005, p. 428).

In the study of emotion and its potential implications on marketing, arousal has been linked to early processes of awareness and attention (Boucsein, 1992), brand recall (Belanche et al, 2014) but also later steps of traditional marketing theory models like information retention and attitude formation (Holbrook et al., 1982).

In a study investigating the very specific effects that arousal has on ad effectiveness, desirability etc. (Aaker et al., 1986) the effects that this thesis is trying to investigate, the impact of the
preceding ad, were also discussed. In this study, Aaker et al. (1986, p. 15) found that “a warm ad (warmth being defined as: ”a positive, mild, volatile emotion involving physiological arousal and precipitated by experiencing directly or vicariously a love, family, or friendship relationship”)
preceded by another warm ad would be less effective than a warm ad preceded by a non-warm ad” and that “the humorous ad generated higher liking, recall, and purchase likelihood when not preceded by a humorous ad - as did the irritating ad when preceded by another irritating ad (presumably because it was then less irritating)”.

The methodology and definitions in terms of valence and arousal proposed by Aaker (1986) are for operational purposes somewhat abstract and ambiguous. Furthermore, while the tools used in Aaker et al’s study are somewhat outdated, the experimental design and approach has very much been an inspiration to this paper as it does focus on very similar aspect of emotion and the effects on ads.

Furthermore, the effects on memory encoding and decoding have also been extensively researched. Studies have shown that ad recognition can be linked to arousal levels, and it has also been found that there is a positive relationship between familiarity and arousal (Jeong et al., 2012).

To contrast these highly relevant findings, it has however also been suggested that highly arousing stimuli can distract subjects from processing of ads, the ability to process central messages and recall material (Singh et al., 1989) as well as “reducing the ability to identify attributes presented in the ad, and increasing the difficulties to encode ad content” (Shapiro et al. 2002).

Belanche et al. (2014) however, researched this potential paradox in ad design and suggested that “high arousal design increases ad attitude, brand recall and purchase probability compared to a neutral arousal design”.

Regardless, there seems to be a clear consensus within the field that arousal has significant impact on how marketing and advertising efforts are perceived - although there seem to be a disagreement on whether the effect is positive or negative.

Having established what arousal is and its importance for marketing and advertising there is an obvious clash with dominant market research in terms of assessing arousal. Because arousal is
automatic, volatile and nonconscious, simply asking people how aroused they are will not suffice. Thus, a discussion of how to better measure arousal will be presented below.

4.2.1 Measuring Arousal

In the pursuit of measuring and assessing arousal, eye tracking was deemed the most suitable method under the present circumstances, mainly because of two reasons; first, there is a significant input lag from methods such as heart rate and galvanic skin response, making it harder to track instantaneous changes in arousal during short time spans. Secondly, because printed advertisements are exclusively experienced through the ocular sense, it seemed like a more intuitive choice to track pupillary changes rather than for example perspiration or heart rate changes.

One of the pioneer studies of effect of emotion on pupil dilatation was Hess who in 1960 found that subject’s pupils dilated when exposed to pleasant pictures while shutting down when exposed to unpleasant pictures (Hess & Polt, 1960). While the result had several methodical flaws and the results have been very hard to replicate since it still was the start of research trying to find correlation between emotional responses and pupil changes.

As understanding of statistics, sound methodological procedures and technological advances have increased over time, more recent studies have been more successful in investigation the connection between emotional arousal and pupil changes.

Steinhauer, Boller, Zubin, and Pearlman (1983) found that subject’s pupils dilated when exposed to both pleasant and unpleasant pictures while Aboyoun & Dabbs (1998) exposed male and female subjects to pictures of dressed and naked people to find that pupil dilation did not differ between genders as arousal was near same levels.
In more recent studies, Bradley et al. (2008) found similar effects as data revealed that “following the initial light reflex, pupillary increases were larger when participants viewed pleasant or unpleasant, compared to neutral pictures”. Figure 6 illustrates this finding. Notice the initial drop in pupil size, which is attributed to the light reflex.

![Figure 6](image)

The effect on pupil dilation from emotional arousal is not restricted to effects from visual stimuli only, but has also been found to exist following auditory stimuli. Partala et al. (2000) was able to find significant differences in pupil dilation when subjects were listening to affectively engaging sounds compared to listening to neutral sounds. This finding suggests that stimuli do not have to be visual for the emotional arousal to affect pupil dilation. However, this paper will focus exclusively on visual stimuli.

In accordance with the notion that the intensity of the stimuli is related to pupil size (regardless of valence), Janisse (1973) claims that: “The most consistent finding relates greater dilation to increases in the intensity of stimulation, whether positive or negative; that is, pupillary activity is linearly related to intensity continuum of affective stimulation, but curvilinear related to the valence (positive/negative) continuum”. The linear relationship is illustrated in Figure 7.
In the context of this study, it is important to remember that measurement of arousal through pupil dilation does not tell anything about the valence of the stimuli creating the arousal. While valence and arousal are the two most commonly used dimension used to understand the nature of emotional information, they are also two dimensions that are not necessarily measured in the same manner. While arousal is defined on a scale from calm to highly excited, valence is commonly linked to behavioral patterns of approach and avoidance behavior.

In that sense, arousal can tell the intensity of the emotional response that is being experienced but not whether that response comes from a stimuli that is of positive or negative valence. As Bradley et al. (2008, p.12) found that: “pupil dilation is determined by emotional arousal, independent of whether pictures are pleasant or unpleasant in hedonic valence”.

Because stimuli with different valence might affect each other, it is important to understand the distinction between how to determine valence and how to determine arousal. This section exclusively focuses on the subject of arousal. The determination of valence will be discussed in the next section.

Another important aspect in understanding pupil dilation is the acknowledgment that other non-emotional effects can cause the pupil to dilate or restrict. Especially cognitive load is positively
associated with pupil dilation (Hyönä. et al., 1995; Kahneman and Beatty, 1966) and luminosity have likewise been proven to affect pupil dilation, which can bias results if not taken into account. These issues together with a more precise account of the actual calculations of the eye tracking data will be provided in a later section on the experimental design.

Finally, other methods of measuring arousal within neuroscience does exist. Some of the most commonly known and used methods are respiration and heart rate, skin conductance (galvanic skin response), EEG and fMRI.

Both fMRI and EEG are popular tools in neuroscientific studies but neither was chosen for this study. First and foremost, fMRI is very expensive and was not readily available for the purpose of this paper. Secondly, fMRI has primarily been used to establish spatial activity in the brain related to certain types of stimuli. While some authors such as Anders et al. (2004) has used fMRI to investigate arousal, it remains a less suitable alternative because of the availability of other, simpler, less expensive and more intuitive methods.

EEG can in many cases also produce very insightful findings by measuring brain activity while recording electrical activity in the brain. However, the data output from EEG measurements is very complex and subject to a wide range of potential biases, and again having the availability of other, simpler and more intuitive methods this method was also deemed unfit within the scope of this theses.

Respiration/heart rate and galvanic skin response has previously been used for several studies on arousal or subjects relating to this concept (Appelhans et al. 2005; Bradley et al. 2008; Groeppel-Klein 2005). While skin conductance and respiration and heart rate has been useful in other findings, it was decided that pupil size would be more suitable due to the previously mentioned input lag.

While a triangulation of methods could have been employed, the scope of this thesis limited such an exercise.
4.3 The role of valence

The other aspect of emotion or affective states is valence. Valence in essence, is the direction of the emotion and typically measures from negative to positive. Valence, as used in psychology, especially in discussing emotions, means the intrinsic attractiveness (positive valence) or aversiveness (negative valence) of an event, object, or situation (Frijda, 1986).

The description of valence has ranged from a broad spectrum of adjectives such as good-bad, optimistic-pessimistic, positive-negative, complete-incomplete and timely-untimely (Mehrabian, 1996).

Like arousal, valence also carry significant marketing implication. Several theories from social psychology have been suggested in relation to this valence effect. It has widely been argued that the valence of people’s current affective state at the time of exposure to an ad will color its evaluation (Forgas, 1995). Valence also have, for example, been linked to affecting consumer judgment in a mood-congruent direction (Clore et al., 1994; Forgas, 1995; Goldberg et al., 1987; Gorn, et al., 1993; Mayer et al., 1992).

Furthermore, Aaker (1986) found that the valence (although labeled ‘warmth’ and not being defined exactly like valence, it still very much resembles the concept) of a preceding commercial seemed to affect the interpretation (and possible effectiveness) of the subsequent commercial. Additionally, tracking the valence of comments online has been used to establish ROI on social media investments (Hoffman & Fodor, 2010).

In relation to the effects of valence on arousal, Libby et al. (1973) found that unpleasant pictures created larger pupil dilation than pleasant. This was concluded by comparing the pupil dilation with cardiac deceleration - something Bradley et al. (2008) was not able to replicate. Bradley et al. (2008, p.606) did however agree with the conclusion that emotional arousing images correlated with pupil dilation as it was found that “During picture viewing, both skin conductance and pupillary changes were greatest for emotional, compared to neutral, pictures and these effects were accentuated for the pictures rated as most highly arousing. Furthermore, when the picture was used as the unit of analysis, the relationship between skin conductance and pupil changes
persisted, even when specific effects due to luminosity were removed using hierarchical multiple regression”

One flaw that might explain this difference between the findings of Bradley et al. (2008) and Libby et al. (1973) was that the latter did not consider level of arousal of the stimuli pictures, a factor the latter group of authors did take into consideration by using IAPS rated stimuli (Lang et al, 2005), which is a database of images that has indexed images in terms of valence and arousal.

While the experimental design of this paper does not rely on such IAPS rated images, it is designed to allow for some categorization of the included ads, so that negative and positive valenced images should be equally arousing. As such, differences in reaction to the ads can potentially be contributed to differences in valence, rather than speculations that it might as well could have been due to differences in arousal level.

In the context of this paper, it is argued that almost all ads are seen in a context of other ads, and that these ads will vary significantly in terms of their intrinsic valence. In a random TV commercial block it’s easy to imagine a car commercial, a laundry detergent commercial, an anti-smoking commercial and an insurance commercial being seen next to each other. Those ads will vary significantly both in terms of valence and arousal. The argument that valence can then influence subsequent evaluation makes this aspect highly relatable to the subject of this paper.

Having defined valence and its importance for marketing, a discussion regarding how to measure the concept will be presented below.

4.3.1 Measuring valence

To assess valence, a self-reported framework was created inspired by Gorn et al. (2001). One could question the rationale in using self-reported scales for assessing valence, but there were several reasons for doing so. First of all, there has been several studies that rely on subjective ratings of valence (Feldman, 1995; Bradley et al., 2008; Aaker, 1986; Lang, 1995), all of which are highly regarded researchers within their fields.
Secondly, valence does not, to the same extent, rely on the same assumptions about non-conscious processes as arousal does, suggesting that conscious reports of valence should be much less problematic than conscious reports of arousal. This is illustrated in Figure 8.

![Diagram](image)

**Figure 8**

As discussed earlier in the section on definitions of emotions and feelings it was illustrated how expressions and autonomic changes happen within seconds and beyond the consciousness of consumers. This can apply to valance as well. For an example it has been proven that valence can guide attention as it has been found that objects that are most emotionally relevant are also more likely to draw subject’s bottom up attention (Genco et al, 2013).

In that sense, it would be problematic to ask for subjective ratings of valence as it according to our earlier definitions, perception of valence can happen within milliseconds (guiding attention) and without conscious awareness. However, self-reported aspects of emotions is also possible meaning that subjects within seconds, minutes, hours or days are able to consciously reflect upon the valence of objects.

To illustrate, when exposed to an image of an half-naked, attractive person, a subject is not likely capable of reporting subtle differences in processes relating to the autonomic nervous system such as perspiration, heart rate frequency and pupil size (all indicators of arousal) - that would be the autonomic changes and expressions that happens within seconds and are unconscious. However,
a subject would likely be capable to report if the image is found pleasant or unpleasant (in essence, valence) - even within a few seconds. Hence, while valance in some regards can be perceived unconsciously, it can also be reflected upon consciously in a way that arousal for example cannot. Therefore, using subjective ratings to define valence is a valid method.

The real obstacle is then to choose what scales to employ when asking subjects to rate the ads. As described, many different interpretations of how to frame valence has been suggested, for example; good-bad, optimistic-pessimistic, positive-negative, complete-incomplete (Mehrabian, 1996).

The framework suggested in this paper draws its inspiration from Gorn et al. (2001) when defining the scales on which valence is measured. However, some issues in terms of how participant interpreted these categories emerged. This issue and how it was dealt with will be elaborated on in the section on variable definitions.

While it is acknowledged that other methods exist, it was deemed that subjective ratings would be an appropriate method of measuring valence in the scope of this paper, as other relatively new and respected researcher has relied on the same methods and the effect of non-conscious processes (that subjective rating cannot capture) should be limited within this parameter.

Other potential tools for measuring valence includes facial recognition software, approach/avoidance from the screen and EEG.

Facial recognition software was incorporated into the eye tracking system used in this experiment, but initial testing showed very unintuitive readings and unreliable results.

Another option was to include approach/avoidance measurement in terms of distance to the screen. However, this methodology is fairly unexplored and with a very limited academic body of research to lean on.

EEG has also been used to measure valence. Studies have found a correlation between what is called hemispheric prefrontal cortex asymmetry and approach-avoidance behavior which should be closely linked to the intrinsic valence of the examined stimuli (Berkman, 2010). However, as already mentioned in the section on measuring arousal, EEG was not found fitting within the scope of this theses.
Clearly, conscious reporting as a research methods has, as extensively argued, its limitations but it was nevertheless found to be the most suitable option.

4.4 Relationship between arousal and valence

As mentioned previously, arousal and valence are two highly related concepts and are usually described as the two main components of emotions (Oatley et al. 2006). Their dynamic interrelation has been the subject of much research.

Following the logic originally presented by Janisse (1974), and repeated in numerous experiments since (Bradley et al. et. al 2008; Steinhauer et. al 1983; Abboyoun & Dabbs, 1998), it will in the context of this thesis be assumed that the relationship between arousal and valence follows a bi-directional, u-shaped curve. This relationship is illustrated in Figure 9 (Coan & Allen, 2007).
This figure illustrates the relationship between the level of stimulus pleasantness (i.e. valence), its behavioral direction (appetitive motivation - approach or defensive motivation - avoidance) and the level of arousal. As can be seen, in general, the more pleasing or displeasing (i.e. positively or negatively valenced) the stimuli are, the stronger the related intensity of the emotion is (i.e. higher levels of arousal). On a casual, intuitive level this relationship seems reasonable. If you feel very positive about something (valence), you are likely to feel stronger about it (arousal), and the same should be applicable for negative stimuli as well.

The relationship can also be displayed as a simplified U-curve, Figure 10. The graph is merely a conceptual illustration of the relationship between valence and arousal rather than an actual, true map. For example, in the illustration it would be possible to identify a completely non-arousing image bearing no valence whatsoever. However, studies has proven that even mundane everyday objects such as teapots carry so called “micro valence” (Lebrecht et al. 2012).

![Figure 10](image)

For the purpose of facilitating the ease of reading, this relationship and its associated implications, it will from this point be referred to as the “U-Curve” or the U-curve theory.
As demonstrated, both arousal and valence are highly relevant to marketing. Having established the difference between the two concepts, different tools to assess and measure them (and the selected methods for this thesis) and the interrelationship between the two concepts, a theoretical summary proposing a framework for the research design will be presented below.

5 Theoretical summary

Having gone through literature and empirical studies in the previous sections, it has been illustrated that traditional marketing research methods might not be sufficient as they often are only capable of captioning conscious processes and are based on consumers ability to reflect upon something they by definition are not aware of. It has also been illustrated that nonconscious processes account for a significant amount of decision-making and can have crucial effects on awareness, attitude, behavior etc.

With the inspiration of priming experiments it has been showed how consumers are able to process subliminal information and be affected by products/brands below the conscious awareness (Dijksterhuis et al., 2005; Fitzsimons et al., 2008; Bargh & Chartrand, 2000).

This effect of unrelated stimuli was further investigated by the likes of Forgas (1995) and Schwartz (2010) and their AIM models/Affect as information investigated how different stimuli (often nonconscious) might affect consumers judgements and evaluations. In the value-based model of choice (Plassmann et al, 2012), the representation and attention stages will affect all subsequent stages of brand/product evaluation but what this model – alongside many others – seem to do is focus mostly on how related stimuli affect each other.

This thesis, on the other hand, poses the question of how different (and unrelated) ads with very different motivational values (negative vs. positive valence) might affect each other – quite possibly beyond the conscious mind of the consumer. Given the limited scope of this thesis, the focus lies mainly on the 2 aspects of arousal and valence based on 3 relevant assumptions:

1. Arousal has, in several studies, been linked to mechanisms highly relevant to consumer behavior. This includes awareness, attention, and retention and attitude formation.
2. Valence is highly differential across different ads and according to several theories this should affect consumers (AIM, affect as information)

3. Arousal and valence have in several studies been proven to correlate.

    With arousal and valence in focus, it is argued that a combination of an eye-tracking study (measuring pupil dilation) and a behavior study (measuring subjective rating) would result in a valuable empirical material aimed at investigating the underlying role of context/sequence.

6 Research design

This chapter will commence with a very brief description of the idea and rationale behind the experimental design aimed at answering the research question of “How are consumers’ nonconscious reactions and conscious judgement of ads affected by the context and sequence of which the ads are presented in?”.

A more detailed description of the 2 groups that will be included will be presented later. Although the design between the two groups have many features in common, they will be described separately in hope of providing a clearer image of the process.

The very basic idea was to compare the emotional arousal (as measured by pupil dilation) and subjective ratings (as a proxy for experienced valence) for a number of different ads between two groups of subjects. One group were shown 36 ads in isolation in a completely randomized order, the other group were shown 15 ads (based on a selection of the same 36 ads) in a fixed, specific order consisting of three blocks with 5 ads in each block, i.e. 15 ads in total. The first group performed the rating task after each individual ad while the second group got to rate all five ads at the end of each block.
While it is claimed that the condition for group 1 is that they viewed the ads in “isolation” and free from context, it is perhaps an impossible task to truly create any experiment completely free from context (unless having each single participant only viewing one isolated ad and then leaving the experiment, which for practical reasons is not possible. Even under that condition it is perhaps not possible to claim that the ad was viewed completely free from context). However, participants in group 1 performed the rating task after each single ad, and because the ads were shown in a different, randomized order for each participant, it can be expected that the average effect on non-conscious reactions and conscious judgements of those ads are much less affected by the “context” than for group 2. There are two main reasons for this is expected; 1) by asking participants to rate each single ad immediately after it is shown, participants are forced to engage in conscious and deliberate processing of that ad which should lessen the spillover effect, and 2) since automatic, emotional reactions occurs within the first few seconds of stimulus exposure, having a rating task (that on average lasts four to seven seconds) in between each ad, should greatly reduce the same spillover effect.

Even if some form of effect from context might still be present in the experimental design of group 1, such effects should be diminished to a point where it is still possible to do meaningful comparisons with group 2 where there is a fixed, specific context that is consistent and identical across all participants. Thus, in the scope of the below presented experimental designs, data and results from group 1 will be treated as the ads had been viewed in isolation.

The objective of the experiment is to investigate the effects of context and sequence on arousal and subjective rating. As this is a three-part experiment, what is being investigated are more specifically 1) the relationship between valence and arousal; 2) the emotional arousal effects that might spill over when viewing advertisement in a context and 3) whether or not any specific valence related, directional effects can be identified when viewing ads in a specific sequence. In the following section hypotheses are defined, research design is presented, and the selected data collection techniques are described.
6.1 Hypothesis definition

Based on the literature review and discussion of theories, combined with the research question of this paper, the following hypotheses are defined:

6.1.1 Correlation between pupil size and subjective valence score

Because of the bidirectional nature of the relationship between arousal and valence, two hypotheses, H1a and H1b were set up:

\[ H1_a: \text{An increase in pupil size will significantly correlate with an increase in subjective rating of positively valenced images (Increase meaning a more positive rating).} \]

\[ H1_{a0}: \text{An increase in pupil size will not significantly correlate with an increase in subjective rating of positively valenced images (Increase meaning a more positive rating).} \]

\[ H1_b: \text{An increase in pupil size will significantly correlate with a decrease in subjective rating of negatively valenced images (decrease meaning a more negative rating).} \]

\[ H1_{b0}: \text{An increase in pupil size will not significantly correlate with an increase in subjective rating of negative valence images (Increase meaning a more negative rating).} \]

6.1.2 Effects of ad-viewing in a context rather than isolated

\[ H2_a: \text{Viewing ads in a context rather than isolated will significantly affect pupil dilation.} \]

\[ H2_{a0}: \text{Viewing ads in a context rather than isolated will have no significant effect on pupil dilation} \]
H2a: Viewing ads in a context rather than isolated will significantly affect subjective ratings of valence.

H2b: Viewing ads in a context rather than isolated will render no significant difference on subjective ratings of valence.

6.1.3 The spillover effects of differently valenced images - sequential effects

H3a: In a fixed position sequence of ads the previous ad will affect the subsequent ad in the direction of the preceding ads’ valence. (For example a negative ad preceding a positive ad will cause the positive ad to be rated less positive).

In order to test the above defined hypotheses, experiments were conducted using both eye-tracking as well as a related rating task asking for the participants subjective ratings of the ads. The goal of these experiments was to investigate the impact of context on both subjective ratings and nonconscious bodily arousal (as measured by pupil dilation).

In reference to the above described hypotheses and the previous explained experimental design (which will be elaborated on for each group in the upcoming sections), the illustration, Figure 11, below serve as a chronological overview of the research process and how hypotheses were tested.

![Figure 11](image_url)
6.2 Operational definition of variables

As suggested in the concept definition, *arousal* and *valence* are the two operational variables for the proposed experiment. A detailed account for how they were assessed and calculated are presented below.

6.2.1 Arousal

As demonstrated, several authors have found convincing evidence that pupil dilation changes are good predictors of emotional arousal. Many different methods for calculating pupil dilation changes has been suggested. In the two experiments conducted to answer the research question of this thesis, pupil dilation was calculated in a way inspired by Janisse (1974) and Bradley et al. (2008). These two studies are some of the most widely cited involving pupil dilation and calculate arousal by taking “… the percentage difference between the pupil size during the stimulus presentation and the average size of the pupil during the presentation of the control slides before and after the stimulus slide”. This is illustrated in Figure 12.

![Figure 12](image-url)
In other words, arousal was calculated by looking at the relative difference between the average pupil size evoked during the six second grey slide and the peak pupil size during stimulus exposure.

The relative increase from baseline (the baseline was calculated as an average size of the pupil during the six second grey slide) to peak pupil size displayed during exposure of target ad would then be determined to be the arousal level of that particular ad. The baseline pupil size of the grey slide should represent the “true size” of the pupil for that specific luminance level because no other factors than luminosity should really impact the pupil size in this specific moment. Both content related emotional arousal and cognitive load should be very limited when viewing an image of random pixels.

Furthermore, other external factors like drugs, pain, drowsiness etc. had been controlled for before initializing the experiments (see experimental design chapter). One could argue that what should have been compared was the baseline (average) of the grey slide with the average pupil size during the exposure of the stimuli (the ad) instead of comparing the baseline pupil size with a peak pupil size during stimulus exposure. It is, however, argued that this would not have been representative. The emotional response from an arousing image is not constant when being exposed to a image for 6 seconds. Partala et al. (2003) showed in a study (also using 6 seconds exposure) that emotional response was likely to be present within the first two seconds after stimulus exposure (starting at about 400 milliseconds after exposure). After the initial reaction, more cognitively driven processes such as reading and analyzing start taking place, thus potentially affecting the pupil size in ways not necessarily strongly related to arousal.

As these effects were not interesting for the purpose of this thesis and as the mathematics behind calculating an average pupil size value would put a larger emphasis on the remaining four seconds (of which emotional reaction should play a smaller part). The peak pupil size was, based on that, deemed to be a better point of reference as the pupil size peak was consistently found within the first two seconds of stimulus exposure.
Partala et. al (2003) states that; “The visual inspection of the timeline curves revealed that the averaged pupil dilations coherently followed a similar type of curve. A sudden increase in pupil size at about 400 ms from the stimulus onset was followed by a slower decrease.”

In other words, the pupillary reaction is not constant throughout the entire exposure time, and by using the mean of the entire exposure time, the arousal effect on pupil size would be greatly diminished.

Instead, by investigating the peak pupil size, we should see when the effect of emotion and or cognitive load is at its highest (please remember that all luminosity effects should have been accounted for by using the normalizing grey slide) and this comparison is therefore much more relevant for the purpose of this thesis.

Definition of arousal, within this thesis is thus - The percentual difference in pupil size between the average pupil size following a six second pre-stimulus grey slide and the peak pupil size evoked during target stimulus (ad).

6.2.2 Valence

As described previously, the valence score was based on subjective ratings where subjects were instructed to rate each image from 1-9 on 3 scales;

- Sad – Happy
- Distressed – Delighted
- Unpleasant – Pleasant

In the initial research design, valence was assessed by calculating the sum of the average rating of each scale for each participant. This number was used as a determination of valence score for each ad. As described, the framework of the scales was taken from Gorn, Pham & Sin (2001). Figure 13 illustrates to basic principle behind this idea.
However, during the pre-testing of the experiment and after examining the results, a potential problem was identified with these 3 scales. Many participants reported that they struggled with deciding on for example how sad or happy an ice cream ad was (but found it easy to report how pleasant or unpleasant it looked) or how distressing or delighting an insurance ad was. It was consistently observed that the most intuitive and easily used scale was the “Pleasant-Unpleasant”.

Furthermore, it was found that several participants in the pre-test expressed difficulties in finding meaningful separation between the scales and having several scales were found to be confusing.

Taking these two issues into careful consideration, it was therefore decided to only include the “Unpleasant-Pleasant” scale as a determinant of valence. This decision will be elaborated on in the later discussion section on validity and what limitations that might have for the paper. Figure 14 serves to illustrate to revised rationale.
Definition of valence, within this thesis is thus - *The average subjective rating on the unpleasant-pleasant scale.*

### 7.1 Group 1 - Experimental design and rationale

A pool of 36 real ads were selected from various sources online, primarily from brandsoftheworld.com. Ads and brands that were likely to have been seen before were excluded from the experiment to avoid potential bias from previous experiences with those ads or brands.

As it has been proven in numerous studies, previous exposure to a stimuli (pictures, signs, faces, etc.) will increase liking of these stimuli, often referred to as the “mere exposure effect” (Zajonc & Markus, 1988). Exposing consumers to the same brand or ad several times will also increase liking and attitude formation of these (Shapiro, 1999). For this reason, ads and brands that were likely to have been seen by participants before were excluded. Furthermore, all participants were upon completion of the experiment asked if they recognized any of the ads or brands.

As the mere exposure effect can happen even when consumers are unaware of the fact that they have previously seen the ad (Zajonc & Markus, 1988), there still is a theoretical risk that some
participants had previously seen some of the ads - and were biased to be more positive towards these. Hence it is recognized that it is impossible to completely rule out some participants had indeed seen certain ads or brands before.

Out of the selected 36 ads, 12 were assumed to be of pleasant/positive valence. These 12 supposedly pleasant images included smiling faces, cute animals, happy families, sexually suggestive ads etc. Another 12 ads were assumed to be relatively neutral in terms of valence consisted of generic ads for insurance, banking services, computer software etc. Lastly, the remaining 12 ads were assumed to be of unpleasant/negative valence as they were mostly graphically explicit, provoking ads such as anti-smoking ads, ads against obesity, animal cruelty ads and suicide prevention ads etc. Illustrations of the ads can be found in appendix 2.

These ads were for the purpose of classification and overview labeled “Positive 1-12”, “Neutral 1-12” and “Negative 1-12”. One might suggest that this name scheme could have provided for some potential source of bias. However, as the participants were never exposed to the actual names of the ads, such bias should not be present. Furthermore, the labeling was not the determining factor in the later analysis and categorization of data. Would an ad, for example, be labeled “Negative 6” and receive a positive valence score, it would be treated as a positive ad with all associated implications.

The intention was that the range and spectrum of the included ads would at least vaguely be reminiscent of the range of ads one could expect to be exposed to in real life during a normal day reading the newspaper, riding the bus, watching television and so on. The ambition was to select roughly the same amount of ads in each category (pleasant, neutral and unpleasant).

Furthermore, ads that contained excessive amounts of text were excluded because it was found in an initial test phase of the pre-test that subjects tended to ignore the emotional content of these ads and focus on the text instead, which was counterproductive to the goal of the intended design.

Another issue with text is that cognitive load has a significant impact on pupil size, which in the case of big variance between the ads, would make it hard to accredit differences in pupil size
between ads to emotional responses (which is the purpose of this paper) rather than cognitive load. Because of these two reasons, the maximum number of characters allowed was set to 100.

Lastly, only ads that were of reasonable size were included, (in this case reasonable size means ads that could likely be fit and shown in a magazine or newspaper).

7.1.1 Manipulation of ads

The only manipulation done to the ads was the conversion to black and white (i.e grey scale) instead of their original colors. As the most important biological function of the pupil is to dynamically react to changes in environmental illumination, the ads were adjusted to grey scale in order to minimize the effect of picture luminosity on pupil dilation (as this effect was not of particular interest) (Bradley et al. 2008).

The iMotions attention tool software includes a grey scale slide option to be displayed before each target stimulus. The grey slide is computed by the software and seeks to match each target picture’s luminosity in order to provide a baseline pupil dilation before the exposure to the stimuli. The software analyses each pixel of the image and scrambles all pixels, creating an “image” that contains exactly the same pixels, but in a “white static” kind of order. This means that the grey slide is of equal luminosity as the original ad but free from any signs, symbols, and meaning or otherwise emotionally affecting content. This allows for readings of the baseline pupil dilation that are used as a comparison with the pupil dilation measured during viewing of the real ad. Figure 15 serves as an example.
Displaying this grey slide before each target ad should therefore adjust for any luminosity effects. However, it has been found that the human eye is differently sensitive to colors in the red, green and blue spectrum which makes it hard to estimate picture luminance for pictures presented in color (ibid). Acknowledging this, in addition to displaying the grey slide before each ad, it was judged most appropriate to display all ads in grey scale as well to avoid extreme results from participants being sensitive to particular colors.

It was acknowledged that manipulating the way the ads are presented (from color to grayscale) provide for a potential source of bias. However, it was judged appropriate to sacrifice some external validity in order to have a more controlled experimental design, as far as pupil dilation goes. The color range and luminosity of the ads was adjusted using Adobe Photoshop CSS 5.6.
7.1.2 Subject sample

For the first experiment, twenty subjects in total were recruited from a business university in the Copenhagen area. Two out of these were excluded due to poor quality of the eye-tracker readings, one was excluded because of interruption during the process and one was excluded for having taken medicine just prior the experiment (medicine that might have altered the subject's pupil size). In total, this resulted in 16 included, valid subjects.

All subjects were right handed male business students with normal to corrected vision. The age ranged between 22 and 30 years. No monetary or other form of compensation was offered for participation. Nonprobability convenience sampling was employed to recruit participants for the first group (Hair et al. 2009). Participants were of Danish, Slovenian, American, Icelandic and German origins.

Females were excluded from the study because it has previously been found that there are some gender-related differences in pupil size related arousal (Wallen & Rupp, 2007). Since some of the calculations are done using mean values across all participants, it would provide a potential source of distortion or bias to the data if both genders were included. For example, sexually suggestive images and ads might be differently arousing between genders. This would potentially have offered some issues as some of the ads were of such character.
As illustrated in the two graphs above (Figure 16), there seems to be gender-related differences in pupillary reactions to differently stimuli of different valence. Thus, Coan & Allen (2007, p.33) concludes that: “The affective space for IAPS pictures for men and women are slightly different, with men showing tighter coupling (higher linear correlation) between pleasure and arousal ratings for pleasant pictures, whereas women show a tighter coupling between pleasure and arousal (higher linear correlation) for unpleasant pictures.”

### 7.1.3 Research equipment, software & location

The first experiment took place at the CDN Research lab at Copenhagen Business School. Upon entering, participants were briefly informed about what was expected of them and that they would be completely anonymous. Participants were seated approximately 63 cm from a computer screen with a resolution of 1900x1200 pixels with a Tobii 60XL eye-tracker attached. The Tobii 60XL eye-tracker operates at a frequency of 60 hertz. This translates into one reading every 16\textsuperscript{th} millisecond. The data from the eye-tracker were recorded and computed through the software “iMotions Attention Tool”. The output from the eye-tracker was analyzed and sorted using Microsoft excel and SAS JMP 11.2 statistics analysis software.

All subjective ratings were sorted and analysed using the same software (Microsoft excel and SAS JMP 11.2) as the data output from the eye-tracker.

### 7.1.4 Other Methodological considerations

To provide for a consistent level of lighting, all blinds were down and the ceiling lights were the only light source apart from the computer screen. After being briefed on how the equipment worked, the eye-tracker was individually calibrated for each participant to provide as good quality of the data as possible.

The experiment commenced with an on-screen slide of information with everything the participants needed to know, such as the duration of the experiment, how the scales were to be
used and instructions for head-position etc. This was intentionally kept in text-format only as verbal instructions might have had a variance in length, tonality and loudness which might have had provided for a source of bias.

Following the info-slide, all participants got to rate two practice ads before starting the experiment, to familiarize themselves with experimental design. This was done in order to minimize confusion and cognitive load during the real experiment. Before proceeding, all participants had a chance to ask any questions they might have had.

All of the 36 ads included in experiment 1 were shown for 6 seconds each with an additional 6 second grey scale slide preceding each image. After viewing each ad for six seconds, a survey page was shown, asking the participants to rate the ad on the three scales “Pleasant-Unpleasant”, “Happy-Sad” and “Delighted-Distressed”. There was no time constraint to complete this rating task and most participants took between four to seven seconds to complete each survey page. By informing the participants that there was no time limit in completing the survey slides, the intention was to lessen some stress for the participants as stress has proven to have a profound effect on pupil dilation (Wang et. al 2009). Although pupil dilation was not measured for analysis during the rating task, (pupil dilation was only recorded during the grey slide and target ad exposure) the stress created from a time limit when rating the images could potentially have spilled over. This could have biased the results and because of this, no such time limit was given.

Below (Figure 17) is an illustration of a small sample of the experiment chronology. Note that this is not the definitive order of ads, but an example from a single participant.
All slides except for the information slides and the two test ads at the start of the experiment were shown in a completely randomized order. The ambition was to minimize potential spill-over effects in regards to subjective rating and pupil dilation. Because the research question of this thesis investigates contextual effects, the first conducted experiment, described above, were to be used as a point of reference for participants emotional reactions when no context was present (see previous discussion regarding free-from-context). Had the ads been displayed in a fixed order, one highly arousing image might have created a spillover effect on the next image, affecting the readings for the second image.
The design was intentionally kept as short as possible to minimize fatigue, boredom and dropping motivation. In group 1, the experiment time ranged between 14 minutes 36 seconds to 21 minutes 55 seconds. After completing the experiment, participants were allowed to ask any questions they might have had regarding the design and purpose of the study.

### 7.2 Group 1 - Results

Group 1’s experiment was designed to serve two purposes for this paper. Firstly based on results from group 1, a number of ads were to be included as comparison with group 2. Secondly, it would serve as a point of reference to group 2 when testing all hypotheses relating to contextual effects (H₂ and H₃). The full results are displayed in Figure 18.

Therefore, the purpose of group 1 would then be to identify the images with the highest valence score and highest arousal (inducing the largest pupil size change from baseline), which should correlate according to previously mentioned U-curve theory. This correlation corresponds to Hypothesis H₁a & H₁b.

<table>
<thead>
<tr>
<th>Image name</th>
<th>Pupil size change</th>
<th>Subjective rating (Pleasant scale)</th>
<th>Categorization</th>
</tr>
</thead>
<tbody>
<tr>
<td>Positive</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Postive 4</td>
<td>2.40%</td>
<td>7.88</td>
<td>Positive</td>
</tr>
<tr>
<td>Postive 8</td>
<td>9.92%</td>
<td>7.56</td>
<td></td>
</tr>
<tr>
<td>Postive 9</td>
<td>1.40%</td>
<td>7.56</td>
<td></td>
</tr>
<tr>
<td>Postive 3</td>
<td>4.11%</td>
<td>7.50</td>
<td></td>
</tr>
<tr>
<td>Postive 10</td>
<td>3.55%</td>
<td>7.31</td>
<td></td>
</tr>
<tr>
<td>Postive 2</td>
<td>1.95%</td>
<td>7.19</td>
<td></td>
</tr>
<tr>
<td>Postive 6</td>
<td>7.53%</td>
<td>7.19</td>
<td></td>
</tr>
<tr>
<td>Postive 12</td>
<td>7.23%</td>
<td>6.81</td>
<td></td>
</tr>
<tr>
<td>Postive 11</td>
<td>8.73%</td>
<td>6.69</td>
<td></td>
</tr>
<tr>
<td>Postive 7</td>
<td>2.97%</td>
<td>6.50</td>
<td></td>
</tr>
<tr>
<td>Postive 5</td>
<td>6.03%</td>
<td>6.13</td>
<td></td>
</tr>
<tr>
<td>Neutral 11</td>
<td>0.52%</td>
<td>6.13</td>
<td>Neutral</td>
</tr>
<tr>
<td>Neutral 1</td>
<td>4.46%</td>
<td>5.88</td>
<td></td>
</tr>
<tr>
<td>Neutral 12</td>
<td>1.11%</td>
<td>5.88</td>
<td></td>
</tr>
<tr>
<td>Neutral 1</td>
<td>0.54%</td>
<td>5.81</td>
<td></td>
</tr>
<tr>
<td>Neutral 5</td>
<td>3.07%</td>
<td>5.69</td>
<td></td>
</tr>
<tr>
<td>Neutral 8</td>
<td>4.69%</td>
<td>5.38</td>
<td></td>
</tr>
<tr>
<td>Neutral 2</td>
<td>3.47%</td>
<td>5.31</td>
<td></td>
</tr>
<tr>
<td>Neutral 10</td>
<td>5.54%</td>
<td>5.31</td>
<td></td>
</tr>
<tr>
<td>Negative 2</td>
<td>5.73%</td>
<td>5.00</td>
<td>Negative</td>
</tr>
<tr>
<td>Neutral 3</td>
<td>2.54%</td>
<td>4.88</td>
<td></td>
</tr>
<tr>
<td>Neutral 4</td>
<td>1.86%</td>
<td>4.88</td>
<td></td>
</tr>
<tr>
<td>Neutral 6</td>
<td>3.22%</td>
<td>4.88</td>
<td></td>
</tr>
<tr>
<td>Neutral 9</td>
<td>6.63%</td>
<td>4.75</td>
<td></td>
</tr>
<tr>
<td>Neutral 7</td>
<td>7.30%</td>
<td>4.63</td>
<td></td>
</tr>
<tr>
<td>Negative 4</td>
<td>6.44%</td>
<td>3.25</td>
<td></td>
</tr>
<tr>
<td>Negative 12</td>
<td>5.67%</td>
<td>2.69</td>
<td></td>
</tr>
<tr>
<td>Negative 5</td>
<td>2.60%</td>
<td>2.63</td>
<td></td>
</tr>
<tr>
<td>Negative 3</td>
<td>5.33%</td>
<td>2.50</td>
<td></td>
</tr>
<tr>
<td>Negative 7</td>
<td>4.45%</td>
<td>2.38</td>
<td></td>
</tr>
<tr>
<td>Negative 10</td>
<td>3.23%</td>
<td>2.31</td>
<td></td>
</tr>
<tr>
<td>Negative 9</td>
<td>4.92%</td>
<td>2.25</td>
<td></td>
</tr>
<tr>
<td>Negative 1</td>
<td>8.93%</td>
<td>2.13</td>
<td></td>
</tr>
<tr>
<td>Negative 11</td>
<td>2.54%</td>
<td>2.13</td>
<td></td>
</tr>
<tr>
<td>Negative 6</td>
<td>2.74%</td>
<td>1.88</td>
<td></td>
</tr>
<tr>
<td>Negative 8</td>
<td>6.35%</td>
<td>1.63</td>
<td></td>
</tr>
</tbody>
</table>

*Figure 18*
7.2.1 Hypothesis H1a

The ads that had a higher average score than 5 (the midpoint of the scale) were classified as “positive ads”. To investigate the correlation between pupil size and rated valence scores for these positive ads, a Bi-variate, linear regression model was employed. The model shows virtually no correlation between the two variables (\(N=16, R^2=0.002\)). Full results are displayed in Figure 19:

While the model displays a slightly positive correlation coefficient, indicating that subjectively rated valence scores for positive images varies in a positive direction together with an increased pupil size, the strength of that relationship is very weak, as displayed by the \(R^2\) value very close to 0.
Virtually no correlation was present between the pupil size and the valence score of any positive ad, thus we are not able to reject the null hypothesis for H1a - “An increase in pupil size will show no significant correlation with an increase in subjective rating of positive valence images (Increase meaning a more positive rating).”.

7.2.2 Hypothesis H1b

The ads that had an average score of less than 5 were classified as “Negative ads”. To investigate the correlation between pupil size and rated valence scores for the negative ads, a Bi-variate regression model was employed. The model shows virtually no correlation between the two variables ($N=16, R^2=0.0082$). Full results are displayed in Figure 20:

![Summary of fit](image)
Similar to H1a, no correlation was present between the pupil size and the valence score of any negative ad. Thus we are not able to reject the null hypothesis for H1b - *An increase in pupil size will not significantly correlate with an increase in subjective rating of negative valence images (Increase meaning a more negative rating).”*

Not being able to reject the two null hypothesis for H1a and H1b, some methodological implications had to be considered. While the lack of correlation, and explanations for this, will be discussed in the results discussion, the method for choosing the images used in group 2 will be explained in the following chapter.

### 7. Selection of ads for group 2

Having tested all 36 ads in the first experiment, a number of these ads were included and arranged in a specific order and context in the second group. As the experimental design for the second group included three blocks of ads, with five ads each, 15 ads (consisting of 7 negatives, 5 positive and 3 neutral) in total were to be selected for comparison. The reason for not choosing 5 ads from each category is described in more detail in the experimental design for group 2. In order to choose which of the 36 images to include in the comparison for the second group, some issues had to be considered.

If, for example, chosen ads from the pool of positive ads had been more arousing than the chosen ads from the negative pool, a potential bias would have been present. Had the, for example, positive ads been significantly more arousing than the negative ads, it would have been hard to determine if the observed effect (should there be an effect) were due to the valence of ads or just from being more arousing. Therefore, the ambition was to select roughly equally arousing ads from both the positive and negative subgroup of ads together with the least arousing images from the neutral group.

The rationale for selecting the least arousing ads from the neutral group was to investigate if these ads were more susceptible to the context related effects than other, more arousing ads.
Selecting equally arousing images were not as straightforward as initially expected. The original experimental design predicted that it would be possible to assess the subject’s level of arousal based only on their pupil dilation. This proved to be very difficult.

The U-curve theory would suggest that for those ads in the group of 36 ads, that are rated most extreme on the valence score scale, a stronger pupil dilation would be expected, (as this is the defined measure of arousal) compared to less extreme rated ads. Certainly, it would be expected that the images that were rated most neutral would also evoke smallest emotional arousal, and thus displaying the smallest effect on subject’s pupil size.

As described in testing of H1a and H1b this was not the case. Taking this finding into account, the criteria for selecting which ads to include in experiment 2 was the following:

By using the U-curve and it’s associated implications, 15 ads were selected to be included in the second experiment based on their valence score - using valence as a proxy of arousal (in accordance with the U-curve theory). Figure 21 is a brief graphical illustration of this rationale.
In this graph, the ads “Negative 10” and “Positive 9” are plotted according to their valence scores as seen on the X-axis. The average valence score of “Negative 10” was 2.3125 and the average valence score of “Positive 9” was 7.5626. As evident, they are both very similar in their distance from being neutral, which would entail a score of 5 (this means that they are similarly strongly valenced but in opposite directions). Following the assumption about the U-shaped relationship of arousal and valence, it can thus be assumed that they are likely to be very similar in their ability to generate arousal.

The negative and positive images were chosen by identifying the images with the highest (most positive) and lowest (most negative) valence scores, with comparable distance from 5 (the neutral score) making them (in theory, according to the u-curve theory) equally arousing but with opposite directions of valence.

The neutral images all had valence scores as close to 5 (being the perfect middle value between positive and negative, hence neutral) making them (in theory, according to the u-curve theory) as non-arousing as possible.

Obviously, it was impossible to identify any images that had the exact same valence score and hence it was not possible to identify any images that would be 100% equally arousing. There was however put great effort into choosing images that were similar and hence this is not deemed a major problem, as the difference was minor between all chosen images. By choosing images that in theory should be equally arousing, the intention was to create a design that were able to attribute differences between the groups (H3) to valence, which would not be possible had all images not been equally arousing.

Finally, while the ads; Neutral 10, 3 and 4 are referred to as “neutral ads” in the experiment, this is not a completely accurate description. Finding an image that is completely unarousing and with no valence either positive or negative is impossible. As a replacement, it was necessary to choose the images the closest to being neutral (and hence non-arousing) as possible.
8 Group 2 - Experimental design and rationale

The experiment with group 1 provided data to calculate mean valence scores for each ad and as described, these data were used to identify some of the most positively rated and some of the most negatively rated images that were found to be equally arousing.

Since the first group was designed to capture the participant’s reactions and ratings in randomized isolation, the experimental circumstances for the second group needed to be designed to do the opposite, i.e. be able to capture the reactions and ratings in a context.

To simulate a context where ads are shown and perceived together rather than in isolation, three blocks of five ads each were created. In contrast to the first group, participants in the second group were asked to rate the five ads, only after viewing the whole block of ads. Figure 22 is an illustration of one such block of ads:
The three blocks of ads were composed accordingly;

**Chronology**

<table>
<thead>
<tr>
<th>Block 1</th>
<th>Block 2</th>
<th>Block 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Negative 9</td>
<td>Positive 9</td>
<td>Neutral 4</td>
</tr>
<tr>
<td>Positive 8</td>
<td>Negative 3</td>
<td></td>
</tr>
<tr>
<td>Negative 10</td>
<td>Neutral 10</td>
<td>Negative 12</td>
</tr>
<tr>
<td>Positive 4</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Negative 7</td>
<td>Negative 8</td>
<td>Positive 10</td>
</tr>
<tr>
<td>Neutral 3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Negative 6</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Figure 23*

The rationale for placement of the differently valenced stimuli within the blocks was the following:

Block 1 and 2 essentially are identical except the negative and positive images have switched place. In other words, image 1 in block 1 and 2 are equally arousing, but they are opposites in terms of valence. The same goes for image 2 in block 1 and 2 and so forth.

By doing this it was possible to investigate if a sequence had a specific effect. For example, since the 3rd image in both block 1 and 2 was an equally non-arousing neutral image it was possible to compare if there was any difference in effect when being preceded by a negative or positive image. Essentially, the design allowed for an investigation of several different scenarios of how differently
valenced (but equally arousing) images affected each other going from negative to positive, positive to negative, etc.

The 3rd block differs in order and sequence from the first 2 blocks. This design was more exploratory by design. Essentially, two highly arousing negative ads were placed together to see if that would create a larger effect on the following positive images, compared to when only being exposed to 1 negative image (like in block 1). This could be called an experiment in seeing if there is some kind of buildup effect when several equally valenced images are placed together and their subsequent effect on different valenced images.

Several other sequences could have been chosen and added to the experiment, however, due to the risk of mental fatigue and even mental overload, it was determined that keeping the experiment reasonably short would likely produce more useable data.

8.1.1 Manipulation of stimuli

Since the same identical ads were used for both of the groups, they were subjected to the same manipulations, i.e. conversion to grayscale.

Again, referring to the problem statement of this paper the idea was to investigate the effects of the context and sequence. By using the exact same images that group 1 had been exposed to, and keeping all other controllable variables constant - thus effectively only changing one aspect between the 2 groups - the context.

8.1.2 Subject sample

For group 2, another sixteen participants in total were recruited from a business university in the Copenhagen area.

All subjects were right handed male business students with normal to corrected vision. The age ranged between 24 and 34 years. No monetary or other form of compensation was offered for
participation. Similar to group 1 in the pre-test, nonprobability convenience sampling was employed to recruit participants for the second group (Hair et al. 2009). Participants were of Danish, Romanian, Norwegian, Icelandic and Swedish origins.

To keep the two groups coherent and comparable, females were also excluded from the second group.

8.1.3 Research equipment, software & location

As the experiment for the second group were conducted in the same facilities using exactly the same equipment, a description of these are surplus. The only present difference was how the images were arranged as explained above.

8.1.4 Other methodological considerations

The experiment commenced with a similar information page (amount of text, font size etc. were kept constant) but with slightly adjusted information as the experimental design differed slightly.

Similar to group 1, participants in group 2 were allowed to do two practice ads to familiarize themselves with the software and design to avoid confusion and excessive cognitive load.

Just like for the first group of participants, each image were shown for six seconds with a six seconds grey slide preceding each image to control for luminosity. However, after viewing the first image for six seconds, the second, third, fourth and fifth image followed (with the accompanying grey slides). During the whole duration, participant’s pupil size were measured. After all five images in the block of ads had been viewed the participants were asked to rate the ads following the same valence framework as group 1.

In group 1, only the pleasant-unpleasant scale was utilized when determining valence for an ad. This also applied to the 2nd group. However, the other two other scales (Happy-Sad and Delighted-Distressed) were kept on the survey slides for one main reason.
Would the scales have been removed from the second group’s survey slides, there would have been a significant difference in the design of the two groups. Since the ambition was to create as equal and similar experimental circumstances as possible to allow for comparisons between the two groups, it was deemed that although the scales were not included in the calculation of valence scores, removing them might have had unforeseen consequences on factors relating to pupil size dilation such as cognitive load. Had group 1 got to rate each ad on three scales, and group 2 got to rate each ad only on one scale, chances are that the accumulative cognitive load would have had an unwanted impact on the data.

8.2 Group 2 - Results

Group 2’s experiment was designed to provide data that primarily would allow for a test of hypothesis $H_2$ and $H_3$.

The same procedure as described in group 1 was used to determine valence score and arousal (pupil size) measurement. While, based on the data and results from group 1, it proved difficult to compare arousal levels (based on pupil size) between different ads within the same group, this problem was not present when comparing the same image across the two groups.

When keeping the same image constant across the both groups, elements such as luminosity or other external factors should pose less problems, as these factors can be assumed to be equally present for both groups. Hence, it should be possible to attribute any differences in pupil size between the groups to the difference between viewing ads in a context rather than isolated. Essentially, the only two variables that was likely to be affected by this change in arrangement should be arousal and cognitive load. Hence, pupil size as a measurement of arousal change was now possible (as a comparison), whereas as it was found in experiment 1 that it might not be the best measurement to solely index arousal level.

Below is a comparison of the results from both group 1 and 2. The tables include Ad names, average pupil size change from baseline, valence rating and the categorization of the ads. While
participants in group 1 got to rate 36 ads in total, only ads that were included as comparison to group 2 are included in the table below (Figure 24).

<table>
<thead>
<tr>
<th>Group 1</th>
<th>Group 2</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Original name</strong></td>
<td><strong>Pupil size change</strong></td>
</tr>
<tr>
<td>Positive 4</td>
<td>2.40%</td>
</tr>
<tr>
<td>Positive 8</td>
<td>9.52%</td>
</tr>
<tr>
<td>Positive 9</td>
<td>1.40%</td>
</tr>
<tr>
<td>Positive 3</td>
<td>4.11%</td>
</tr>
<tr>
<td>Positive 10</td>
<td>3.55%</td>
</tr>
<tr>
<td>Neutral 10</td>
<td>5.54%</td>
</tr>
<tr>
<td>Neutral 3</td>
<td>2.54%</td>
</tr>
<tr>
<td>Neutral 4</td>
<td>1.86%</td>
</tr>
<tr>
<td>Negative 12</td>
<td>5.67%</td>
</tr>
<tr>
<td>Negative 3</td>
<td>5.33%</td>
</tr>
<tr>
<td>Negative 7</td>
<td>4.45%</td>
</tr>
<tr>
<td>Negative 10</td>
<td>3.23%</td>
</tr>
<tr>
<td>Negative 9</td>
<td>4.92%</td>
</tr>
<tr>
<td>Negative 6</td>
<td>2.74%</td>
</tr>
<tr>
<td>Negative 8</td>
<td>6.35%</td>
</tr>
</tbody>
</table>
Since contextual effects were being tested, an important point is that the first ad of each block (Negative 7, 9 and positive 4) was not exposed to a context - as they had no preceding image to influence them.

As expected, none of these images revealed any significant changes in pupil dilation compared to group 1 or difference in (as will be discussed in H2b) valence score. As a result in, reality 7 out of 12 images had significant changes in pupil size (since the remaining 3 could not be affected by context). From this point on any reference to contextual effects of the experiments will be made on the comparison of 12 images rather than 15, as only 12 images could realistically have been affected by the context.

In other words, while 15 ads were part of the experimental design and testing of H2 and H3 only 12 ads where subject to contextual effects. Had the 3 ads starting each block displayed any significant changes in pupil size or valence score it could be argued that some external factor had not been controlled for, but this was not the case.

Figure 25 displays which ads were affected, their average pupil size dilation from baseline, the statistical significance and the direction of change.

<table>
<thead>
<tr>
<th>Ad name</th>
<th>Group 1 Average pupil size increase from baseline</th>
<th>Group 2 Average pupil size increase from baseline</th>
<th>Prob &gt;</th>
<th>Significant</th>
<th>Direction of change</th>
</tr>
</thead>
<tbody>
<tr>
<td>Negative 3</td>
<td>5.329%</td>
<td>7.286%</td>
<td>0.0301</td>
<td>*</td>
<td>↑</td>
</tr>
<tr>
<td>Negative 6</td>
<td>2.736%</td>
<td>3.982%</td>
<td>0.4461</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Negative 7</td>
<td>4.452%</td>
<td>6.317%</td>
<td>0.4343</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Negative 8</td>
<td>6.347%</td>
<td>11.776%</td>
<td>0.0081</td>
<td>*</td>
<td>↑</td>
</tr>
<tr>
<td>Negative 9</td>
<td>4.919%</td>
<td>5.232%</td>
<td>0.8863</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Negative 10</td>
<td>3.323%</td>
<td>6.410%</td>
<td>0.0903</td>
<td>*</td>
<td>↑</td>
</tr>
<tr>
<td>Negative 12</td>
<td>5.671%</td>
<td>9.607%</td>
<td>0.0685</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Positive 3</td>
<td>4.112%</td>
<td>2.398%</td>
<td>0.3039</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Positive 4</td>
<td>2.405%</td>
<td>3.634%</td>
<td>0.6174</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Positive 8</td>
<td>9.916%</td>
<td>15.359%</td>
<td>0.0535</td>
<td>*</td>
<td>↑</td>
</tr>
<tr>
<td>Positive 9</td>
<td>1.396%</td>
<td>4.542%</td>
<td>0.0399</td>
<td>*</td>
<td>↑</td>
</tr>
<tr>
<td>Positive 10</td>
<td>3.553%</td>
<td>5.779%</td>
<td>0.3798</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Neutral 3</td>
<td>2.542%</td>
<td>5.654%</td>
<td>0.0726</td>
<td>*</td>
<td>↑</td>
</tr>
<tr>
<td>Neutral 4</td>
<td>1.860%</td>
<td>7.145%</td>
<td>0.0033</td>
<td>*</td>
<td>↑</td>
</tr>
<tr>
<td>Neutral 10</td>
<td>5.543%</td>
<td>8.206%</td>
<td>0.2087</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Figure 25*
The direction of change column indicates the change from group 1 to group 2. As can be seen, for all ads that differed significantly between the groups, the difference was positive, meaning that the average pupil size increase from baseline for those ads were larger in group 2 than in group 1.

8.2.2 Hypothesis $H_{2b}$

$H_{2b}$: Viewing ads in a context rather than isolated will significantly affect subjective ratings of valence.

$H_{2b0}$: Viewing ads in a context rather than isolated will render no significant difference on subjective ratings of valence

In order to test hypothesis $H_{2b}$, a single independent sample, two-tailed t-test ($N=16, p<0.1$) was conducted for each of the compared ads. The null hypothesis $H_{2b0}$ was rejected for 3 out of 12 images, meaning that there was a significant difference in how those 3 ads were rated on the Pleasant-Unpleasant scale between the two groups.

Figure 26 displays which ads were affected, their valence score, the statistical significance and the direction of change.

The direction of change column indicates the change from group 1 to group 2. As can be seen, for all positive ads that differed significantly between the groups, the difference was negative, meaning that the average rated valence score was significantly lower in group 2 for these three images.
8.2.3 Hypothesis H3

*H3: In a fixed position sequence of ads, the previous ad will affect the subsequent ad in the direction of the preceding ads’ valence. (For example a negative ad preceding a positive ad will cause the positive ad to be rated less positive).*

This hypothesis was tested by examining those pictures that were found to differ significantly between the two groups in terms of valence (see H2c results). Should this hypothesis be true, a positive ad preceded by either a negative or neutral ad could be expected to be rated less positive. A neutral ad preceded by either a positive or negative ad would be expected to be either less favorably or more favorably rated depending on the valence of the preceding ad. A negative would be expected to be rated more positive should it be preceded by either a positive or neutral ad.

Because of the fact that the ads in group 1 were shown in a randomized order, only data from group 2 was included when performing this analysis.
The positive ads found to differ significantly between the groups were Positive 3, 9 and 10. All three of these ads were rated significantly lower in group 2 as compared to group 1. By examining the sequence of ads displayed, it can be seen that all of these affected ads are preceded by lower rated ads, thus supporting the hypothesis H₃.

Figure 27 displays the methodology behind H₃.
9 Quality evaluation of experiment

Before discussing the results and their implications, a brief quality evaluation discussion will be presented. Describing any issues regarding validity or reliability will allow for a more meaningful discussion of results, as any potential limitations will have been clarified.

9.1 Validity

Validity is the extent to which a research method, concept or conclusion represents the real world accurately (Zikmund et al., 2010). In order to control for this in the experiment of this paper, the following factors were considered.

Initially, the idea was to allow the participants to judge the ads on a continuous, sliding scale. However, due to technical limitations in the software, this proved to provide too much bias as the starting position of the slider could not be defined, meaning that the slider would always start at the lowest possible value.

This would represent a significant problem according to the anchoring effect as several studies have shown that subjects are highly influenced by numbers put into a context where evaluation is needed (Kahneman, 2011, p.118): “Any number you are asked to consider as a possible solution to a problem, that requires an estimation, will create an anchor effect”.

In the context of this paper, having the slider start at the lowest number could potentially make participants rate the images lower on average than had the slider started at another number. Regardless of what number the slider would have started on, there would have been a potential source of bias as this could have introduced the above mentioned anchor effect.

Because of this, instead a 9-point Likert-style fixed point scale with no set starting number was chosen. The decision to deviate from the standard 7-point Likert scale and employ 9-point was
grounded in an observation during pre-testing. Many subjects expressed that they felt constrained by the 7-point scales. By allowing subjects a greater level of freedom, the ambition was to more precisely capture their subjective feeling towards the ads.

9.2 Experiment environment

The main idea of the experiment was to investigate the effects of a context rather than viewing ads in isolation. In order to identify any effects of such a context all other variables had to be controlled for, which is why the experiment was conducted in an artificial controlled laboratory environment.

Several external factors were adjusted for including the notable luminosity effects that can disturb readings of the pupil size. Not only was the lighting in the room kept constant for all participants, images were also adjust to grayscale and finally, the two groups were compared across percentage differences in changes, rather than individual pupil size (which differs from person to person).

Despite these precautions, while a laboratory experiment has a higher internal validity, it is acknowledged that it, in most cases, has a limited external validity (Hair, Bush, & Ortinau, 2009). Field experiments can create a higher level of realism, as they are conducted in a more natural setting. In regards to print ad viewing, several factors were set up (necessary to create concrete comparable results) that might not translate well to reality. For example, participants had 6 seconds to view each ad, which would differ considerably in a real life context.

Additionally, other senses were excluded, so that sounds, scents or touch were not part of the design. Sensory information from these would likely have created noise in the data, and as pupil dilation measurement is very sensitive, the experiment design called for a strict control of external factors. In real life, the paper quality of the magazine, the smell of the traffic or the sound of the radio might have far reaching implications outside the scope of this thesis.

Conducting an experiment with conditions more similar to an real-life environment (in terms of experiment location, competing sensory information, e.t.c.) could potentially be an interesting direction for future research. It is however noted that the current methodology of eye tracking might not be the best suited option for such an experiment.
9.3 Confidence level

Normal practice in most academic papers that are relying on different statistical testing of hypotheses, is to use a confidence level of $\alpha=0.05$. However, testing the hypotheses defined in the two presented experiments above, a confidence level of $\alpha=0.1$ is instead used. This might seem like a too generous acceptance of type-1 kinds of error. However, it was found that several ads were very close to being significant on a $\alpha=0.05$ level. As all ads had been altered considerably by making them greyscale, it was argued that the ads would have been significantly less arousing than if they had been presented in their original colored appearance. While this greyscale adjustment was necessary to control for luminance effects, it was also acknowledged that any effects on arousal related pupil size dilation might have been limited by the very design of the experiment. Thus, had this modification not been performed, more extreme results would likely have been observed. Based on this, it was decided that a significance level of $\alpha=0.1$ could be justified.

9.4 Participant bias

A potential bias among experiment subjects are when clues or hints of the experiments objective is revealed and subjects alter (not necessarily consciously) their answers to what they think is the expected or right one (Zikmund et al., 2010).

A great deal of effort was put into the design of the experiment so that no participant could realistically guess what was being tested. Furthermore, in the debriefing phase all participants were asked if they had any guess to what they were being tested for and to state that purpose if they had a guess. All participants in some form felt they were being asked to rate the different ads (which, of course was true in some sense) but not a single participant revealed any suspicion that the experiment, in reality, held the context and sequence as the main variables in focus.

Another well-known bias is referred to as the “Hawhorne effect” (Zikmund & Babin 2009). This effect is when subjects act or responds differently because they know they are part of an experiment. This problem should not have been an issue with measurement of pupil size as this is an autonomic and unconscious response that cannot consciously be controlled at will.
Interaction with the researcher was also held at a minimum (during the experiment there was no interaction between the subject and the researcher – all information was giving in writing on a screen), which means there should not have been any reaction during the experiment or any feedback that might have mistakenly have been given.

Still, there is a risk participant altered their ratings of some images (consciously or unconsciously) to adjust to the task they were being tested for (even though no one successfully guessed what they were being tested for).

9.5 Validity of rating scales and adjustment of these

In pre-testing, several subjects gave the feedback that several of the positive images (a positive image would here be defined as a ad with a valence score of 5,1 or above) was hard to rate on the scales and relate to each other. For an example “Positive 8”, and ad consisting of a woman dressed in bikini was rated relatively positive (7,56 rating) among all participants on the scale of unpleasant - pleasant. However, many participants struggled to apply the same positive score to the two other scales (sad-happy and distressed-delighted). Subjects reported that while finding the image pleasant, they didn’t necessarily find it happy or delighting and was confused about whether they should judge if the woman looked happy or if they, themselves felt happy while looking at the woman.

In the negative images (images rated 4,9 or below) this problem did not exist to the same extent. Essentially, subjects had an easier time applying most negative words to negative images (sad, distressed, unpleasant) than they had applying positive words to positive images (happy, delighted, pleasant). An image of a ad against child labor with a picture of child missing fingers (negative 8) would most often be rated low (or negatively) on all 3 scales while a positive image of a woman in bikini (positive 8) would often only be rated significantly positive on 1 scale (pleasant scale).

This presented a problem as negative images, on average, was rated more negatively, than the positives was rated positive – because 2 out of 3 scales didn’t seem to be compatible with the subjects idea of negative vs. positive valence.
It is instead argued that the difference in scores across the 3 scales is not due to a lack of arousal but because the framing of the scales did not apply similarly to ads in both directions (negative and positive).

The pleasant-unpleasant scale was essentially the only one of the 3 scales where subjects understanding of the concept seemed to be equal for both categories.

It was therefore decided to use only the unpleasant-pleasant scales. It does however highlight a potential problem with using subjective ratings when defining valence as the subjective understanding of different words in relation to advertisement does not seem to be congruent across positive and negative images.

9.5 Representativeness of selected images

There were 3 main potential biases with the selection of which ads to use.

First of all, the initial 36 ads chosen were picked using the researcher’s subjective guess of what would be considered negative, positive and neutral (non-arousing). The ads were tested to identify how arousing they were and what valance they held, but despite this it is very possible (and likely or even certain) that other ads exist that would have been more arousing. Using more arousing images would likely have created larger effects.

Secondly, an important aspect of the experimental design was to find equally arousing ads with opposite ratings of valence (positive and negative). As discussed earlier, no correlation between pupil size and valence could be found, following H1a and H1b. Because of this, and with the aid of the U-curve theory, it was necessary to identify which images to include for comparison based on valence score and hence no absolute, accurate value of each ad’s arousal level was established.

Finally, real ads were chosen for the experiment. While controlling for bias by making sure no participant had been exposed to the ads or brands before, other variables could still be relevant like, product category for example. While this could present a potential bias and disturb the results of H3, it was argued to be important to enhance external validity.
Choosing noncommercial images (landscapes, people, animals, etc.) could have eliminated any difference in effect depending for example personal relevance in regards to products category, but it was not found very relevant for the research question of this paper to compare noncommercial images.

9.6 Sample population representativeness

Because voluntary, convenience sampling was employed, there is a potential risk of sample population bias. Furthermore, only including males recruited from an university in Copenhagen further limits the possibility to draw far reaching conclusions about the larger population in general.

9.7 Reliability

Reliability is the goal of obtaining dependability, consistency and reliability over time. The measurement of the experiment should create the same result on repeated trials (Saunders, Lewis, and Thornhill, 2009).

In the two experiments of this thesis, the method was built under strictly controlled conditions and with a highly standardized procedure, with as little human interaction as possible with any subjects in the pursuit of an adequate level of reliability.

However, there are several internal factors that could disturb the result and/or create different results in a repeated trial. For example psychological circumstances and internal states like hunger, thirst or different affective states and moods could alter the results considerable. (Zikmund and Babin, 2009).

Additionally, there is an inherent paradox in that some of the theories that this thesis is testing, might actually influence the result so that results may alter in a retrial. For example the theory of emotional congruence would predict that there is a spillover from ad to ad, because subjects mood
would be altered by an e.g. negative ad and project those negative moods onto the following e.g. positive ad. However, that theory would also predict that the mood any participants enters the experiment with, would alter the perception of the ads being viewed. That means that a retrial could in theory produce a different result depending on the mood of the subjects when they enter the experiment.

9.8 Sensitivity

As previously discussed, the subjective rating of valence was recorded using a 9-point-likert scale. That creates a need for a brief discussion about sensitivity, as “...sensitivity of a scale is important measurement concept, when changes in attitudes are under investigation” (Zikmund et al., 2010, p.309).

Sensitivity refers to the ability of the scale to represent variability. One could argue that the 9-point Likert scale might lack sensitivity to some degree. As noted, the scales were changed from a 7-point Likert scale to a 9-point because subjects, in an initial pretest, expressed a lack of freedom to choose. Still, a 9-point likert scale likely won’t be creating the complete level of freedom that would be ideal. Hence, some participant’s answers might have been slightly different had they had more freedom to choose whatever value of score they wished.
10 Discussion of results

The main objective of this thesis was related to the contextual effects that might exist when viewing ads. In relation to this, specific sequential effects and the relationship between arousal, as measured through pupil dilation, and valence was also explored. Having accounted for the results of the experiment, the following section will now discuss those results and what academic and marketing related implications they might have.

10.1 Discussion of H₁a and H₁b

The testing of hypothesis H₁a and H₁b revealed no correlation between pupil size measures and valence scores. This contradicts the relationship suggested by Janisse (1973). However, before reaching the conclusion that the relationship between arousal and valence does not exists it must be discussed that the lack of correlation could be due to several different issues. Four main reasons are identified to explain why the relationship was not observed in testing of H₁.

1. The U-curve theory does not hold true - There is in fact no such relationship between emotional arousal and stimulus valence. While the U-curved relationship between arousal and valence is intuitive and easy to grasp, and is often a taken-for-granted relationship in social psychology research, there is of course the risk that all studies that has observed the relationship has been faulty. There is also the risk that the relationship is only valid under certain circumstances that was not present under the current experiment.

2. Validity of arousal measurement - Pupil size is not a valid measurement of arousal. If pupil size is not correlated with emotional arousal, the arousal-valence relationship should thus not be observable by measuring the subject’s pupil dilation. However, the relation between the pupil size and emotional arousal has been researched and confirmed numerous times (Janisse, 1973; Bradley et al., 2008; Partala & Surakka, 2003), and there should be no reason to doubt the link between pupil size and arousal.
3. Validity of valence measurement - *The employed rating scales does not capture valence accurately*. As discussed previously, valence is elusive and hard to define, and there is a risk that the scale used in this experiment are not able to fully account for the whole concept. The framework of the scales has however been used successfully previously by authors such as Gorn, Pham & Sin (2001).

4. Distortion from external factors - *Other factors apart from emotional arousal has too much of an impact on pupil size to allow for an accurate assessment of arousal through only measuring pupil dilation*. Considering the amount of different variables that affects pupil size (picture luminosity, cognitive load, anticipation, stress e.t.c.) there is a potential risk that such effects “drown” the effect from emotional arousal and makes it hard to observe in isolation. The relationship with valence could in this scenario be very hard to find. In this case, it would thus be impossible to create some form of index of arousal for each image, but it would however allow for comparison between groups as the affecting variables (picture luminosity e.t.c) is kept constant between the groups.

The first (the described relationship between arousal and valence does not exist), second (eye-tracking lacks validity for assessing arousal) and third potential reason (the valence scale does not measure valence) has been observed and confirmed in numerous experiments, and these should not be the main cause for lack of correlation between the arousal and valence.

Instead, it is argued that it is the fourth condition that is responsible for the lack of observed correlation between the subjectively rated valence scores and the related pupil dilation (an important distinction from emotional arousal) for that ad. The distinction between claiming that there is a lack of correlation between the valence scores and emotional arousal (which we still assume exist) and a lack of correlation the valence scores and pupil size is important because there would be a whole other set of related implications, should the first condition be true.

Explicitly stated, it is assumed that the lack of correlation between pupil size (arousal) and valence scores is due to external factors other than arousal having an impact in pupil size and not because the relationship between emotional arousal and valence does not exist. Understanding what these
external disturbing factors are is now necessary. The main factors affecting pupil size (besides emotional arousal) are:

**Luminance**

The biological function of the pupil to adjust to lighting has a far greater impact on pupil size than any other process, and thus could easily drown out any other effects.

This issue was, as previously mentioned, addressed by having all images converted to grey scale and small corrections were done in order to ensure that luminance was relatively stable across the sample of images. Unfortunately we did not have any software or technical knowledge of any method that would enable us to make all images identical in terms of luminance. To further limit the problem of luminosity, the eye tracking software employed, allowed for a “grey slide” to be shown before all images.

**Mental workload**

Mental workload and pupil size has been proven to correlate in several studies. In one study the pupil size increase was twice as big when subjects calculated 16 times 23 versus when they calculated 7 times 8 (Hess and Polt, 1964).

This tendency has been shown several times as studies show higher mental efforts equal higher pupil dilation. In one study Kahnemann (2011, p. 35) gave participants different mathematical tasks and watched the pupil change live on a screen. He explains how “we had fun impressing our guests with our ability to predict when the participants had given up in regards to the task given... While watching from the hallway we sometimes surprised the participants and our guest by asking “Why did you stop working just now?”, to which the participants would answer “How did you know that?”

To limit the mental workload, a criteria for all chosen advertisement was relatively little text (with a limit of 100 characters) as described earlier.
Drowsiness and fatigue

The effect of drowsiness and fatigue on pupil size is not yet conclusive as some researchers have found that pupil diameter decreases as a function of drowsiness (Lowenstein et al., 1962) while others (Beatty, 1982) have found no change in pupil size. To limit any effect that might exist, the experimental design was on purpose kept relatively short. The length of both experiments were kept as short as possible, and while subjects had to perform task in regards to rating all images, this was not an overly complicated task.

Diabetes, age, pain & drugs.

Patients with diabetes, on average, have smaller pupils although the cause is still not completely uncovered (Cahill et al, 2001). However, no participant reported suffering from this disease.

Pupil size can vary depending on age (Bitsios et al., 1996), but as the sample group was consisted of a fairly similar aged group for both experimental conditions, age-related differences in pupil size should be insignificant.

Pain has also been proven to increase pupil size as pain intensity increases (Chapman et al, 1999). To avoid this problem all respondents were asked if they had any medical issues that would or did cause them any pain. None of the participants reported any immediate or disturbing pain.

A number of drugs (both legal and illegal) can cause the pupil to dilate (Langham, 1965). All participants were asked if they were on or had recently taken any medication. One respondent did in fact report having had a headache and eating medicine to counteract this shortly before being recruited. As we did not have any empirical studies investigating that specific drug and its effect on pupil size, that respondents data was deemed invalid.

All of the above mentioned factors were controlled for but some are harder to limit in their effects than others. In the case of this experiment, mental workload was identified as the most probable factor disturbing the readings.
In testing of H1, it was attempted to label and rank each image compared to each other in terms of arousal. Here, external factors seem to have too big an impact on pupil size, meaning the numbers are distorted and not necessarily representative of arousal. In the specific case of this thesis, mental workload presumably had a big impact on the pupil size. While the task given was relative easy, subjects were given a relative short amount of time to view each image (something that was needed to control that subjects did not view each ad for different amount of periods). This could potentially create some cognitive load as participants had to “decode” the ads and their messages, make up their mind about them and decide how they should rate the ad, all in a relatively short amount of time. While all ads had relatively little amount of text, the variance in amount of text of different ads could also contribute to this problem.

This findings of H1 could imply that using pupil size as an indicator of emotional arousal can be difficult when comparing different stimuli with each other. These might alternate too much and be impacted by other external factors - especially cognitive load and luminosity. However, it is argued that both luminosity and cognitive load was relatively well controlled for in this experiment.

As a logical extension, it could be suggested that eye tracking can have some significant limitations. Eye tracking as a measurement of arousal through pupil dilation should be a valid tool in some cases, but can fall short in others. When comparing identical images across different groups, and controlling for as many external factors as possible, (like in the testing of H2 and H3) differences in pupil size could likely be attributed to arousal.

Thus, it is suggested that eye tracking is better utilized when comparing the same stimuli across different groups where external factors can be controlled for and manipulated with.
10.2 Discussion of H<sub>2a</sub>

H<sub>2a</sub> tested if: *Viewing ads in a context rather than isolated will significantly affect pupil dilation.*

This was tested by performing independent sample, two-tailed t-tests (*N*=16, *p*<0.1), comparing mean values for pupil size change from baseline of each individual ad between the two groups. It was found that for 7 out of 12 images the average change in pupil dilation from baseline differed significantly between group 1 and group 2. The affected ads were negative 8, 10, 12, positive 8, 9 and neutral 3 and 4.

The research question of this thesis includes focus on how consumers’ nonconscious reactions to ads are affected by the context of which the ads are presented in. The findings of hypothesis H<sub>2a</sub> supports the idea that those emotional reactions are significantly affected by the context of which the ad is placed in. Close to 60% of the included ads evoked pupil size changes from baseline in group 2 that significantly differed from group 1 - all of which were *increases* in pupil size. The fact that all affected images differed in the same direction (being more arousing) is notable. This finding would seem to suggest that in contexts with more extreme (arousing) stimuli being present, other stimuli is potentially also being perceived more arousing.

As reviewed in the theory chapter, the difference in evoked arousal could potentially be due to *emotional congruent* interpretations of the ads (Bower, 1981). Following this theory, the suggestion would then be that if participants were placed under a certain mood by seeing some of the images, other images might have been interpreted more congruent with that mood. However, as hypothesis H<sub>2a</sub> investigates the effect on the non-conscious component of emotion, *arousal*, it seems less likely that conscious interpretations should modulate instant, automatic bodily reactions to such an extent displayed in these results.

Another potential explanation could be that being exposed to ads in a context exercise some nonconscious influence on the stage of *attention and representation* as described in the value based model of choice (Plassman et al. 2012). In this case, the participants would non-consciously be more prone to automatically attend to, and notice features in the ads that could serve to alter the degree of arousal that the ad evokes.
If the explanation of the effect can be attributed to an alternation of attention and representation according to the value based model of choice, then there is a fairly significant indication that associative priming can indeed happen between two unrelated ads. By being exposed to a highly arousing ad, it could thus be suggested that the associative network related to such a bodily arousal, would be more readily available when being exposed to the subsequent ad, leading to a more arousing sensation for that ad as well; All of which happens without conscious awareness.

These findings also seem suggest that some of the previously reviewed hierarchy-of-effects models might have an overly simplified and uncritical description of attitude formation. The common three stages present in most models following this logic is a think-feel-do kind of sequence. However, because it was consistently observed that emotional arousal most often seemed to peak within the first two seconds of ad exposure, one could question if the cognitive assessment of the ad really is the first thing to happen and if this assessment really guides our feelings towards the object. The findings in these experiments seem to suggest that cognitive assessment, instead, occurs much later in the process.

For example, as discussed earlier, the Steiner-Lavidge model (DePelsmacker, 2010) proposes that the consumer goes through the stages of Awareness, knowledge, liking, preference, conviction and purchase. However, it fails to recognize that awareness is often modulated by arousal and autonomic bodily responses (Lane et al, 1999), that preference can guide attention, that knowledge is certainly not required for purchase and that liking, preference and purchase often is in contradiction to each other (Genco et al. 2013).

Following the findings of H2a, the notion that awareness and knowledge is always required to form some kind of liking seems like an exaggerated claim. If “liking” (liberally interpreted as valence) can be affected by stimuli that is completely unrelated to the object in focus, such as a context for an advert, and if this processes happens completely outside the awareness of the consumer, then one could hardly claim knowledge is a prerequisite for attitude formation.
While 7 out of 12 significantly affected ads might seem like a big proportion, some of the manipulations done to control for external factors might in fact have limited the impact on emotional arousal. For example, some ads containing blood, skin, bruises, smiles, food etc. could be argued to be less arousing in its grayscale form. Although, no research on this topic has been done to our knowledge, it is perhaps not an impossible suggestion to think that blood is more appalling when colored red as opposed to grey, that naked skin is more pleasing in true skin color, or that an ad for a delicious ice-cream is more appealing when one can see the red colored cherry with the dark colored chocolate sauce running down the sides - an example is suggested in Figure 28:

If the suggestion that natural, colored images is more arousing than pale, grey scale images is true, then perhaps even more images would have been significantly different if this manipulation had not been performed. This also illustrates the rationale for accepting a significance level of $\alpha=0.1$.

These findings combined seem to suggest that influences from feelings and emotions risen from unrelated stimuli, does in fact affect non-conscious reactions to other stimuli. Should the context of which and ad is viewed in be completely insignificant for how that ad is reacted to, the above
result seems highly unlikely. Judging from reviewed literature and theories, there are several potential reasons as to why there is such an effect. The highlighted suggestion for explaining the effects were the theory of emotional congruence (Bower, 1981) and a spillover effect from ad to ad on the representation and attention stage of the value-based model of choice by Plassmann et al (2012), potentially explained by a associative priming effect between ads seen in a context.

10.3 Discussion of H$_{2b}$

H$_{2b}$ tested if: Viewing ads in a context rather than isolated will significantly affect subjective ratings of valence.

This was tested by performing independent sample, two-tailed t-tests ($N=16$, $p<0.1$), comparing mean values for the pleasant-unpleasant scale of each individual ad between the two groups. It was found that 3 out of 12 images had significantly different valence ratings in group 2 compared to group 1. The affected ads were Positive 3, 9 and 10.

Interestingly, all three contextually affected ads were 1) of positive character and 2) rated worse in group 2.

Although most ads did not differ between group 1 and group 2 on the conscious, subjective valence scale, examining just the positive ads, 3 out of 4 ads were indeed affected (1 positive ad was excluded because it was the first ad in the block and thus could not have been affected by context). Judging by this limited experiment, and working under the discussed assumptions, it seems as if the positive ads are more susceptible to contextual factors, leading to altered judgement while the negative ads were more resistant to external influence.

As discussed previously, the emotional congruence theory states that current moods exercise influence on the way that information is being attended to, encoded and remembered (Bower, 1981). Relating this to the value based model of choice (Plassman et al. 2012), one could expect that the first level, representation and attention would be affected by context. If subjects were influenced to more likely automatically attend to negative features (or if the positive features were
attended to less) in the positive ads, a lower subjective rating, like the one displayed for three out of five positive ads, could be expected.

The fact that 7 ads showed significant differences between the groups for arousal (pupil dilation) but only 3 ads differed significantly when comparing subjective ratings of valence is an interesting discussion. Had the present thesis been based on traditional market research assumptions, it would have been assumed that participants were able to consciously report their true feelings and reactions to the ads. What happened on the non-conscious level would have been treated as a peripheral by-product because the subjects would have been assumed to be able to reach and report all relevant information. The emphasis would in that case have been put on differences between the two groups in subjectively rated items, such as the pleasant-unpleasant scale. Slightly exaggerated, the focal question asked by the market researcher would then be - “Well, what did they think of the ads?”.

Thus, concluding that 9 out of 12 images were rated “the same” between the two groups, the logical extension, following these traditional market research assumptions, would be to conclude that context and sequence does not really matter, or only to a very limited extent.

However, as it was just illustrated in the previous section, the differences found in arousal levels between the two groups seem to suggest a contrasting picture. Following the logic of the U-curve theory and because it is now known that arousal affects a large number of processes crucial for consumer behavior such as motivation, affection, attention, memory, liking and behavioral reactions (Belanche et al, 2014; Groepel-Klein, 2005; Boucsein, 1992; Holbrook et al., 1982; Jeong et al, 2012; Shapiro et al, 2002), this gap in results between H2a and H2b instead serves to illustrate one of the main points of this thesis - Alternative tools and methods might in many cases be required for conducting more meaningful and insightful market research.

Hypothesis H2a, (tested with biometric tools as a part of a more neuroscientific toolbox) that cannot be reported consciously, suggest that context matters. Hypothesis H2b tested with more traditionally tool and consciously reported methods, suggests that it might not.

Another potential reason for observing this contradictory gap could be that the validity of the chosen method for measuring valence might be low (see validity discussion in quality evaluation
of experiment). In that case, one could expect more similar results between the two hypotheses if valence could be more correctly captured and measured.

10.4 Discussion of H₃

H₃ states: *In a fixed position sequence of ads, the previous ad will affect the subsequent ad in the direction of the preceding ads’ valence.*

Testing of this hypothesis was done by manually examining the three ads that significantly differed in terms of valence scores between the two groups. Because only a small number of ads were significantly affected, the results should be interpreted cautiously.

All three ads that did in fact differ between group 1 and group 2 were positive ads (specifically positive 3, 9 and 10). All of these were rated significantly worse in group 2 and all of these images were preceded by lower rated images (all ads that had been labeled negative), thus supporting H3. Figure 29 illustrates the concept of valence congruent effects on subsequent ad.

**Negative ad preceding positive ad**

![Figure 29](image-url)
This is, interestingly, as also mentioned in discussion of $H_{2a}$ and $H_{2b}$, very much in line with the emotional congruency theory. Based on the findings of $H_3$ and the emotional congruency theory, it could be suggested that strongly arousing images influence consumer’s mood (in the direction of that ads valence) and that this mood will affect judgment of subsequent ads.

The *Affect-as-information* theory would in this scenario suggest that negative ads possess a greater ability to covertly linger in the mind of the consumer, nudging the consumer towards a more negative mood, affecting the conscious judgement of the positive ads in a negative direction. When consciously or nonconsciously asking “*How do I feel about it?*” the *affect-as-information* theory propose that the consumer will incorporate current feeling into the conscious evaluation of the target. Not always being aware of the true source of their feelings, these feelings from external, unrelated sources might be mistakenly interpreted as originating from the object in focus.

Thus, following the results displayed in $H_3$, it seems as the negative ads affected participants in a more covert, hidden way. If subjects were able to draw the conclusion that “*I feel good right now because I just smiled at the cute kitten in the last ad*” then the influence from the positive ads on negative ads should be limited. However, if subjects had a harder time realizing that “*I feel unpleasant right now because of the abused child I just saw*”, those feelings would be more likely to exert influence on the conscious evaluation of the positive ads.

Following the logic of the *AIM*, for a *heuristic processing strategy*, these findings would suggest that the positive ads to some extent require more cognitive resources as most of the positive ads were affected by context, while no negative ads saw any influence on how they were judged.

Again, bearing in mind that the two conducted experiments in this thesis carry many limitations, one could perhaps speculate that positive stimuli are more subjective and personal, thus requiring more conscious interpretation (*what do I think of that ice cream? the chocolate sauce looks good, but then again, I don't really like cherries*). It is perhaps not unreasonable to speculate that negative stimuli such as blood, wounds and hurt children are, as a result from evolution (in line with how Darwin (1872) discusses emotion), more universally, unconditionally and automatically perceived as negative, thus requiring less mental effort to judge. Should this be the case, one could expect the kind of results displayed $H_{2b}$. 
Because no neutral or negative ads were found to differ significantly between group 1 and group 2 in terms of valence scores, it is a much more uncertain and speculative exercise to examine the influence of the preceding ads valence. Although not statistically significant, it is still interesting to look at the direction of change for all ads rating and compare these changes to the valence of the preceding ad. Keeping in mind that the difference is only significant for 3 out of 12 images, it's worthy of noting that for all neutral ads, the direction of valence still seems to support H3 as neutral 3 and 4 (rated higher in group 2) were preceded by higher rated ads, and neutral 10 (rated worse in group 2) were preceded by a lower rated ad.

Even the single positive ad (Positive 8) that did not change in a statically significant way, were found to confirm the direction of H3. This means that the direction of change for all neutral and positive ads would confirm H3, although only 3 out of those 7 ads were found to have statistically significant changes.

This is contrasted by the negative ads, which in 0 out of 5 cases support H3 (2 of the 7 negative ads are not included in that conclusion as they were the first ad in their respective block and hence could not be affected by context). All changes in the negative ads, although not statistically significant, changed in opposite direction of what H3 would predict. For example Positive 8 preceded Negative 3 but in group 2 the rating of the negative ad is more negative.

Explaining this paradox is difficult, and again referring to the lack of statistical significance, it could just be due to random variables. However, relating to the earlier discussion about how negative stimuli might be more resistant to external influence one, could perhaps consider an evolutionary perspective of the human brain.

While many positive stimuli and their motivational aspects could be seen as crucial for survival of the species in the longer term, such as reproducing (finding persons the opposite sex attractive) and ensuring survival through satisfying nutritional needs (perceiving energy dense food delicious), while negative stimuli such as blood, wounds and other immediate dangers could perhaps, for the purpose of survival in the shorter term, be considered of more acute and crucial character.
Although a highly speculative claim, aversiveness towards certain negative stimuli might thus be more resistant to external influence because of its importance for survival.

This notion is confirmed by Kahnemann (2011, p. 274) who states;

“*When directly compared or weighted against each other, losses loom larger than gains. This asymmetry between the power of positive and negative expectations or experiences has an evolutionary history. Organisms that treat threats as more urgent than opportunities have a better chance to survive and reproduce.*”

To put it simply, it is possible that reactions to negative and positive stimuli work under different conditions in the human brain. While positive and neutral ads seem to be colored by a preceding ads valence, it seems that, judging from our findings, the negative ads are, if anything, reinforced and strengthened in the evaluation of their valence.
11 General discussion & Marketing Implications

One of the major theories inspiring the research question of this thesis was the concept of priming. The notion that advertising has the function of creating motivational priming and hence affecting subject’s goal and from that influencing strongly on behavior was put into context, when investigating how ads then affect each other.

In regards to the finding that 7 out of 12 ads experienced significantly higher pupil dilation and 3 out of 12 ads experienced significantly difference valence scores when comparing the two groups, the findings seems to suggest that ads can indeed affect each other.

While the effect from a preceding ad on a subsequent ad can be both positive and negative, it was not possible to determine any consistent pattern within the scope of this thesis. However, the findings can serve as a strong indicator that ads can have associative priming effects on subsequent ads, which then can affect its motivational priming abilities (and hence the subsequent ads effectiveness).

The presented findings would also then suggest that the notion of analyzing and interpreting ads effectiveness in isolation is outdated and in many cases needs to be reconsidered. No ad operates in isolation and in context they seem to be affected by other ads.

To answer the research question of this thesis: “How are consumers’ non-conscious reactions and conscious judgement of ads affected by the context and sequence of which the ads are presented in?” three main points can be made.

1. Non-conscious, emotional reactions to ads are affected to a large extent by the context of which the ad is viewed in. Spillover effects from highly arousing ads affects the overall level of arousal for other ads and should thus have an impact on ad effectiveness. When placed in a more arousing context, an ad should enjoy (or suffer, depending on desired outcomes for that particular ad) a higher level of arousal as compared to if the same ad was placed in a less arousing context. Even ads that are in themselves relatively neutral should be affected by this phenomenon.
2. Conscious judgements of ads are affected to a lesser extent than non-conscious reactions. Positive ads seem to be more sensitive to external factors and thus seem to be affected to a greater extent than neutral or negative ads.

3. The specific sequence of ads only seem to affect ads of positive valence. Several positive ads that were preceded by negative ads were rated worse, while no negative ads preceded by a positive ad displayed the same kind of dynamics (i.e. rated better because they were preceded by positive ads). Several suggestions to why this was observed has been offered.

So, what do these findings mean in the broader perspective?

For market research, these findings further confirm what has been found by several different authors previously - consumers are seldom able to fully account for, understand and consistently report their preferences. Both arousal and valence was argued to have several marketing implications. When measuring the two concepts, significant differences were found in levels of arousal (a difference in close to 60% of the cases) while much less significant differences were found in terms of level of rated valence (25% of the cases). This is despite the fact that the two concepts have been proven and is argued to correlate (although the experiment of this thesis struggled to find that correlation, likely due to external factors).

Referring to the traditional perspectives of consumer’s research, most of the traditional methods, in various degrees, lean on self-reporting methods. This thesis serves to illustrate the need for incorporating other methods to gain more insightful data. Like the later perspectives in marketing research, after the interpretive turn, it is recognized that consumers might not be rational in their behavior.

However, in the new paradigm discussed previously, it is also recognized that asking consumers might then not make sense (like in consumer research for example, where consumers are argued to be emotional and irrational, yet still employing in-depth interviews as a primary method for data collection).
Relating the findings of H3, to the more traditional marketing models, some weakness of these has been highlighted. The hierarchy-of-effect models downplays or ignores the effect of unconscious emotion and. As has been described, this is a significant problem as a majority of ads in the experiment of this paper were indeed affected in terms of arousal when viewed in a context.

One of the most well-known models in advertising is the AIDA, but as demonstrated, the representation and attention according to the value-based model of choice (corresponding to the 1st A of the AIDA) might be affected by external stimulus. Hence, only leaning on the AIDA framework for evaluating ad and communication effects, one is likely to neglect contextual effects.

According to the FCB and the think-feel vs. high-low involvement dimension framework, Vaughn argues: "for certain products such as sugar, mineral water, paper towels, soap and banks, cognitive elements are important, while for products such as cakes, ice cream and perfume, affective elements seem to have more impact on the buying decision process" (DePelsmacker et al, 2010). In other words Vaughn argues that the effects of affective elements will depend on the product category.

However, while product category was not a main emphasis of the analysis in this paper, it was found that arousal levels (affective elements) were significantly affected for ads that included a diverse set of different product categories including pain relief medicine, anti child-abuse, anti-smoking, milk, diapers, chairs and software technology. Certainly, many of these ads belong to drastically different product categories (based on the think-feel and high/low involvement framework), which according to the FCB grid would make them be differently sensitive to affective elements. However, these findings would suggest that affective elements can be equally important for products/brands/ads regardless of category and assumed level of consumer involvement.

The value-based model of choice was earlier highlighted as an insightful model on branding and communication effects as it incorporated unconscious effect and processes. While the framework of this model still is argued to be highly relevant, the findings in this thesis could suggest an elaboration of the model as it does focus solely on a single brand/stimuli and does not take into consideration the context. By adding a section called "representation and attention of external
stimuli” it was attempted to illustrate how this aspect (in the scope this thesis, external stimuli would specifically be print advertisement) can alter representation and attention of the target brand/ad.

Figure 30 is an illustration of the proposed modified value based model of choice with the additional suggested “representation & attention of external stimuli” incorporated. Please note that the influence from the added representation & attention of external stimuli on experienced and remembered value is faded in the revised model because the conducted research within this thesis does not allow for any conclusions regarding the impact in these later stages of the model.

![Diagram](image_url)

**Figure 30**

While predicted value was not specially investigated, the fact that some conscious rating of positive ads decreased in group 2 compared to group 1, it could be interpreted as an indicator that
this step might also have been affected. It is however not possible to draw any decisive conclusions without further research on the matter.

Finally, based on the design of research conducted in the scope of this thesis, it was not possible to conclude anything regarding impact of external stimuli on experienced and remembered value. However, this would seem to be an interesting path to pursue with future research.

In the revised and suggested framework, the creation of attitude and learning towards ads and brand not only incorporates the unconscious aspect that many traditional model have missed, but also deals with the aspect that spill over from how the context might affect one or several of the value creation aspects of the original model.

Paradoxically, while arguing for the use of neuroscientific tools to be utilized more in consumer research, a finding of the hypothesis testing in H1 was that eye tracking might have some limitations in terms of measuring arousal. While still being argued as a solid tool when comparing the same stimuli under different conditions, it was found to be problematic when trying to compare different stimuli with each other. Simply put, external noise seemed to create too much distortion, and this was even in spite of having controlled for external factors extensively.

Implication wise this could suggest that eye tracking should either be used in combined with other methods, or that other different methods entirely should be used when trying to compare different stimuli and determine their arousal levels relative to each other.

In terms of advertising, these findings seem to suggest that there is a good reason to consider the type of context your advert is going to be placed in. While the exact dynamics of how differently valenced ads affect each other cannot be decisively concluded within the scope of this thesis, there were some interesting signs suggesting that especially positive ads are susceptible to influence from negative ads.

To illustrate a real life implication, this would have significant relevance for example a luxury car marketer. If an ad for a luxury car is placed after a highly negative ad for anti-smoking, the findings of this thesis does to some extent suggest that the car ad would be perceived less positive. That
would very likely impact the effectiveness of that ad in a negative way, and hence the relevance of where in a context you are placed should be very important.

Even when acknowledging that only 3 out of 12 images supported H3, the findings of H2a still suggest that context has very significant effects on arousal - which should affect ad effectiveness.

Thus, while this thesis might not be able to accurately predict exactly how ads are affected by their context, there is evidence that ads are definitely affected by their context - A claim that have often been ignored since many traditional marketing and advertising models focus on analyzing ads and communication in isolation.

The findings presented in this thesis could potentially carry implications for many other marketing practices as well. One direction for further research might be to include other sensory information as variables, paired with other marketing practices such as co-branding, product design considerations and product placement.

The noisy context of modern media platforms are also likely to be susceptible to effects from unrelated stimuli. For example, investigating how moving images such as TV-commercials and more interactive marketing platforms such as modern social media could provide valuable insights into how consumers are affected by context when engaging directly with brands.

In relation to branding effects, it is noteworthy that this thesis has deliberately excluded previously known brands. Conducting the same experiments with known brands that might have personal relevance could potentially reveal another dynamic than the one found in this thesis. An interesting direction for further research could be to investigate if “strong” brands, for example, are more resistant to the influence of unrelated stimuli.
12 Conclusion

This thesis has attempted to answer the research question of:

*How are consumers’ non-conscious reactions and conscious judgment of ads affected by the context and sequence of which the ads are presented in?*

The answer to the research question is that consumers *are* affected non-consciously and (to some degree) consciously by the sequence or context in which they view ads. On an unconscious level, consumers seem to experience a kind of spillover effect from ads in a context, making them non-consciously react to ads more in line with how arousing the context they are displayed in are. On a conscious level, the relative lack of significant differences between the two groups, could be interpreted differently depending on whether you reference the AIM, affect-as-information or the theory of emotional congruence, while it definitely highlights the gap between consumer’s conscious minds and what never reaches consciousness.

To answer the research question, 3 hypotheses were set up and tested.

As a main point of motivation for the thesis was the suspicion that traditional research methods might be inadequate in creating results that give insight into consumers’ minds, an eye tracking study was set up. In order to evaluate ads compared to each other, a hypothesis was set up investigating the relationship between valence and arousal. This relationship had been found to exist in several other studies. This relationship was investigated using pupil size as an indicator of arousal. Somewhat surprisingly, this relationship was not confirmed, although, it was deemed that this was not due to a lack of correlation between arousal and valence, but rather shortcomings in using eye tracking solely as an indicator of arousal. Other variables were argued to have too big of an impact – in the scope of the design within this thesis, mainly cognitive load was suspected of disturbing the results.

Having found this, it was argued that determining arousal through pupil size measurement, should either A) be done by comparing the same stimuli across different conditions (as was ultimately done in this thesis), B) be used in collaboration with other methods or C) be done a much more simple experimental design where external factors can be kept at an absolute minimum.
In testing the contextual effects through $H_{2a}$ and $H_{2b}$ interesting and significant effects were found. In $H_{2a}$ 7 out of 12 ads proved to create significantly higher arousal levels in the second group that had been exposed to ads in a context. Discussing this finding, it was suggested that perhaps a highly arousing context can alter the attention and representation of subsequent ads according to the value based model of choice.

Likewise, later steps of the same model (predicted value, experienced value, memory & learning) might be affected as a result from this, however the research design of this thesis does not allow for any speculations regarding this dynamic.

In contrast to $H_{2a}$, investigating $H_{2b}$ only revealed that 3 out of 12 images differed significantly in terms of rated valence. While this number is harder to draw any conclusions from, it does provide strong argumentation for the need of neuroscientific tools in market research – a main point of the thesis. Since arousal (based on non-conscious measurements) and valence (based on conscious measurements) differed significantly in terms of the measured effect from group 1 (viewing ads in isolation) and group 2 (viewing ads in context), it serves as an illustration that consumers might not be able to report what they are thinking, feeling or even doing.

Finally, having found significant differences when viewing ads in a context it was investigated whether images “colored” the subsequent ad in a valence-oriented direction. Although less conclusive than the other two hypothesis, some indications were found supporting this idea and the effect could indeed be observed, however mainly for positive ads. Positive ads were consistently rated less positive when following right after a negative ad.

Marketing implications for the findings of the thesis would include marketing managers being forced to start considering the context in which to place an ad, as well as understanding a need for utilizing other methods than less dependent on self-reporting.

To illustrate and conceptualize the findings of contextual effects the value based model of choice was elaborated on, including influences from external (and, at least from a strictly rational point of view) irrelevant stimuli.
List of references


Lang, R. M., Bierig, M., Devereux, R. B., Flachskampf, F. A., Foster, E., Pellikka, P. A., & Stewart, W. J. (2005). Recommendations for chamber quantification: a report from the American Society of Echocardiography’s Guidelines and Standards Committee and the Chamber Quantification Writing Group, developed in conjunction with the European Association of Echocardiography, a branch of the European Society of Cardiology. *Journal of the American Society of Echocardiography*, 18(12), 1440-1463.


Appendix 1

Appendix 2

Negative 1 – 6 (From left to right)
Negative 7 – 12 (From left to right)

Every month a child dies by not wearing seatbelt in the back

Animal cruelty abuses live in many ways

Words can kill. Trevor with the message ‘Dye’

It's called suicide because it's your decision
Neutral 1 – 6 (From left to right)
Neutral 7 – 12 (From left to right)
Positive 1 – 6 (From left to right)