Designing for Mood

Affecting Online Judgments by Using a Mood-Lifting Website Design

Master's Thesis

By Caroline Hove Engelheim

Copenhagen Business School 2016

E-business, Cand.merc.(it.)

Supervisor: Mari-Klara Stein

Spring 2016

Pages: 79

Characters: (176.805)

Hand-in date: 14.03.2016

Table of contents

1. Abstract ... 3

2. Introduction ... 3

3. Research Design ... 5

3.1. Paradigmatic Underpinnings ... 6

3.2. Research Purpose and Data Analysis ... 7

3.3. Research Strategy: Case Study of Imerco ... 8

4. Theoretical Underpinnings ... 9

4.1. Our Reasoning Depends on How We Feel ... 9

4.2. How Mood Affects Judgment ... 10

4.3. Mood and Processing Style ... 15

4.4. Mood and Persuasion ... 16

4.5. Mood Induction Methods ... 17

5. Qualitative Methods ... 20

5.1. Expert Interviews ... 20

5.2. Individual Interviews ... 22

5.3. Focus Groups ... 25

5.4. Limitations of the Qualitative Study ... 26

6. Qualitative Findings ... 27

6.1. Mood Cues ... 27

6.2. Individual and Contextual Differences ... 37

6.3. A Mood-optimised Design ... 38

7. Discussion and Implementation of the Qualitative Findings ... 41

7.1. Photos ... 41

7.2. Background Colour ... 43

7.3. Fluency ... 43

7.4. Gift 46
7.5. Advantages 46
8. Quantitative Methodology 47
8.1. Design 47
8.2. Procedure 49
8.3. Instruments 50
8.4. Data Types 52
8.5. The Questionnaire 52
8.6. Hypothesis Testing 53
8.7. The Sample 56
8.8. Limitations of the Experiment 59
8.9. Choice of Statistical Tests 62
8.10. Preliminary Analysis 64
9. Quantitative Results 65
9.1. Descriptive Statistics 65
9.2. Causal Relationships 66
9.3. Correlations 69
10. Discussion of the Quantitative results 71
10.1. The Effect of the Design on Mood 72
10.2. The Relationship between Mood and Online Shopping Related Judgments 72
10.3. The Effect of the Design on Online Shopping Related Judgments 73
11. Conclusion and General Discussion 75
11.1. Summary of Results 75
11.2. Theoretical Contributions and Limitations 76
11.2. Managerial Implications 77
11.3. Ethical Reflections 79
12. Bibliography 80

1. Abstract

With the hardware dealer Imerco's webshop as a case, the present research used a mixed methods approach to examine how online users' mood is influenced by website design and whether their mood affects their online shopping related judgments. Firstly, it was explored how the users' mood is influenced by various website elements. Five website elements that influence the users' mood positively were identified and integrated into a 'mood-optimised' mock-up design of Imerco's website. Secondly, a two-group experiment was conducted to test whether the mood-optimised design affects the users' purchase intention, product and website evaluation, and whether this effect is mediated by mood. As expected, it was found that the mood-optimised design makes the users employ a more heuristic processing style, which indicates that the design has a positive effect on mood. Furthermore, it was confirmed that users' mood affects their online purchase intention and evaluation of a website positively. These findings have great managerial implications, since they suggest that sales can be raised by affecting users' mood by website design.

2. Introduction

Keywords: *Mood, mood induction, evaluation, purchase intention, persuasion, processing style, fluency, website design, e-commerce, online consumer behavior.*

In the 1970s-90s a large number of academic studies established that there is a positive relationship between mood and judgment (Clore, Gasper, & Garvin, 2001; Forgas, 1995; Oatley, Keltner, & Jenkins, 2006). When this theory was applied to a business context, it was found that there is a positive relationship between mood and product evaluation (Gorn, Goldberg, & Basu, 1993; Isen, Shalker, Clark, & Karp, 1978) and purchase intention (Donovan & Rossiter, 1982; Swinyard, 1993). These findings have critical managerial implications, since they suggest that sales can be raised by affecting customers' mood, e.g. through store atmosphere (Donovan & Rossiter, 1982; Spies, Hesse, & Loesch, 1997).

Nevertheless, the retail sector has drastically changed since the 90s: Today a large part of consumer shopping takes place online. The number of online customers and their average amount of spending has grown expansively over the last 20 years and continues to grow (Laudon & Traver, 2015). Businesses therefore have to adapt their sales strategies to the online environment. This makes it relevant to know whether customers' decisions are affected by their mood when their shopping takes place *online*. Few studies have addressed this issue. Likewise, it is relevant to know *how* to affect online customers' mood in an online setting. The only medium for affecting online customers is the website, since the physical

setting cannot be controlled. During this study, it has not been possible to come across any study that explores the various types of website elements[1] that influence mood.

The present research sheds light on these issues by (1) exploring how various website elements influence online customers' (henceforth 'users') mood and (2) seeking to confirm that their mood affects their purchase intention and evaluation of the website and its products. The research questions are:

How can users' mood be influenced by website design? Will improving their mood lead to an increased purchase intention, product evaluation and website design evaluation?

To answer the research questions, the present research conducted a mixed methods study using the hardware dealer Imerco's webshop as a case. Though this research was firmly rooted in a literature review, which provided insights into theory on the connection between mood and judgment, it began with a qualitative study that aimed to explore how various website elements influence users' mood. This study consisted of two expert interviews (with a User Experience Designer and a Conversion Rate Optimisation Manager), five individual interviews and two focus group interviews (with website users). The interviews made it possible to discover five website elements[2] that influence users' mood positively: (1) Photos with certain characteristics[3], (2) a calm and light background colour, (3) fluency of information, (4) a gift and (5) clear information about service and product advantages. These were named 'mood cues' because they signal to the users that the situation is benign. The cues were integrated into a new mock-up design of Imerco's webshop (henceforth the 'mood-optimised design'), which was tested in a two-group experiment. This made it possible to measure the effect of the design on purchase intention, product evaluation and website design evaluation, and to test whether this effect is mediated by mood. Processing style was also measured, because researchers have found that people in a good mood use a more heuristic processing style[4], which makes them easier to persuade with weak arguments (Forgas, 1995; Mackie & Worth, 1989; Schwarz, 2011).

The results suggest that the mood-optimised design makes users employ a more heuristic processing style, which indicates that the design also has an effect on mood. In addition, they suggest that mood

[1] In this report, a website element is defined as any visible element on a website – e.g. photos, colours, text, shapes, etc.

[2] The five website elements are explained in detail in section 6.

[3] The photos have to be calm, neat and unstaged. They should be accompanied by text and display the products in a nice, familiar, and decorative 'dream setting'. This is explained in detail in section 6.1.1.

[4] The heuristic and systematic processing style represent two distinct systems of judgment. Heuristic processing is fast and superficial, whereas systematic processing involves intensive reasoning. Section 4.3. explains this in detail.

significantly accounts for variations in website design evaluation and specific purchase intention[5]. Even though the remaining relationships were not confirmed, they cannot be rejected, since the level of statistical power was 0.20, which means that there is an 80% risk that it was not possible to detect the relationships even though they do exist. The suggestion that website design affects users' mood and thereby their online judgments has critical implications for online marketers, since it means they have to design their website to affect users' mood positively.

In the next section, the research design will be described. This is followed by an explanation of the theoretical underpinnings along with the hypotheses that predict relationships between mood and online judgments (henceforth the 'quantitatively derived hypotheses'). The last part of this section presents the research on methods for influencing mood, which drove the initial coding template for the qualitative analysis. Next, the qualitative methodology and results are described and discussed. This lead to the presentation of a number of propositions, which are merged into a hypothesis (henceforth the 'qualitatively derived hypothesis') that predict how the design will affect mood in the quantitative study. Subsequently, the quantitative methodology and results are described and discussed. Finally, all results are summed up, followed by a discussion of theoretical and managerial implications and ethical reflections.

3. Research Design

In the present research the qualitative and quantitative parts are equally dominant and supplement each other by providing different insights that would not be gained by using only one of them (Harwell, 2011).

The first research question calls for an exploratory qualitative approach because the aim is to discover new data and understand various website elements from the users' point of view. This provides a rich, in-depth understanding of how users understand and affectively react on various website elements and why (Miles & Huberman, 1994). The second research question calls for a quantitative confirmatory approach because the aim is to test the qualitative findings and test whether theory on mood and judgment also applies in an online context (Miles & Huberman, 1994; Saunders, Lewis, & Thornhill, 2009). While the qualitative study explores users' mood as an elastic concept, the quantitative study objectifies mood and measures it on a quantifiable scale. By use of a highly structured methodology, relationships between variables are tested to confirm the hypotheses.

[5] In this report, specific purchase intention is defined as purchase intention towards a specific product. Section 8.3.3. explains this in detail.

Qualitative and quantitative methods are traditionally prescribed by the two contradicting philosophical paradigms interpretivism and positivism. Nevertheless, the present research follows a pragmatist philosophy, which regards the use of mixed methods as beneficial. This will be explained in the following section.

3.1. Paradigmatic Underpinnings

The qualitative part of this research suggests an interpretivist philosophy[6], since the aim is to understand the users in their specific context, whereas the quantitative part suggests a positivist philosophy[7], since it focuses on testing observable entities to build generalised conclusions. This confirms the pragmatist view, which follows the principle that it is possible to work with variations in epistemology[8], ontology[9] and axiology[10] and that the most important consideration for the research is not the philosophy, but the research questions, which will be best answered by use of both qualitative and quantitative methods (Saunders et al., 2009).

The qualitative study takes a subjectivist epistemological stand in regarding both subjective meanings and social phenomena as acceptable knowledge (ibid.). In addition, reality is regarded as socially constructed and hereby socially changeable, and users are regarded as constructed ideas that can only be understood from an intersubjective point of view. This suggests an idealist ontology, which holds that social phenomena are created from the perceptions and consequent actions of those social actors concerned with their existence (ibid.). The qualitative results are regarded as value bound and inseparable from subject and researcher bias, which means that the axiology is also subjective (ibid.). On the other hand, the aim of the quantitative study is to reduce bias. It takes an objectivist epistemological stand in regarding observable phenomena as credible data and in viewing the users as objective entities instead of subjective beings. Mood, product evaluation, purchase intention and website evaluation are thus regarded as objective, observable facts that have a separate existence to that of the researcher (ibid.). This suggests a realist ontology, since users are regarded as real and measurable objects that exist in reality external to social actors concerned with their existence (ibid.). Even though the researcher is mainly regarded as

[6] Interpretivist philosophy emphasises the difference between conducting research among people rather than objects and advocates that it is necessary for the researcher to understand differences between humans in our role as social actors (Saunders et al., 2009).

[7] Positivist philosophy regards social reality as objective and observable entities, uses theory to develop hypotheses and aims for law-like generalisations (Saunders et al., 2009).

[8] Epistemology concerns what constitutes acceptable knowledge (Saunders et al., 2009).

[9] Ontology concerns the researcher's view of the nature of reality (Saunders et al., 2009).

[10] Axiology concerns the role that the researcher's own values play in the research process (Saunders et al., 2009).

objective and independent of the data, which suggests an objective axiology, the stance of this research is that even though validity and reliability issues can be lowered to a minimum they are impossible to avoid. There is always a risk of some respondents not understanding the questions in the way as intended by the researcher or that uncontrollable factors (such as a respondent having a difficult day) will create variance between the experimental groups (cf. section 8.8).

The qualitative findings about user's affective reactions and the quantitative findings about the relationship between design and other variables are context dependent, since they are based on the study of Imerco's website and its users. This means that they are not necessarily generalisable to other webshops. On the other hand, the findings about the relationships between mood and other variables are grounded in theory-driven hypothesis testing and can be generalised to other e-commerce websites as well.

3.2. Research Purpose and Data Analysis

In accordance with the pragmatic philosophy, the research was conducted with both exploratory, explanatory and confirmatory goals, and the analysis was performed both in- and deductively.

The quantitative study was carried out with a confirmatory goal to confirm hypotheses based on findings of previous researchers and an explanatory goal by seeking to explain causal relationships between variables (David & Sutton, 2007). A deductive approach was used for analysing the data, since all variables were derived from theory on mood and judgment (Saunders et al., 2009).

In the individual interviews, an exploratory approach was used by asking open questions without seeking to confirm or reject theories. In the expert interviews, a predominantly exploratory approach was used, since most questions were open, but the aim of the last few questions was to confirm or reject whether the elements that have an effect on mood in other contexts (identified in literature) also influence mood in the present context. The focus groups were conducted with a partly confirmatory and partly exploratory purpose, since the informants[11] were only asked to discuss the elements that had been identified or confirmed in the other interviews. This made it possible to go more into depth with understanding the meaning of these elements.

An abductive approach was used for analysing the qualitative data, which is a mixed in- and deductive approach suitable for mixed methods (Harwell, 2011). A template analysis (a flexible method) was used

[11] The term 'informants' is used to refer to the people who took part in the individual and focus group interviews.

by commencing with an initial coding template based on theory on elements that affect mood in other contexts. Continuously, as the empirical data revealed new categories, these were added to the template, and if no empirical findings confirmed the existing categories, these were removed (illustrated in Figure 1). Thereby, the conceptual framework only represented the key themes and relationships found in the qualitative data, which makes the approach mostly inductive (Saunders et al., 2009). In the coding scheme, categories were structured in a hierarchy with the various website elements in the top and the different aspects of insights related to those elements as the subcategories, going into more depth of analysis for each level in the hierarchy. The final coding tree can be found in the appendix (Appx. 8).

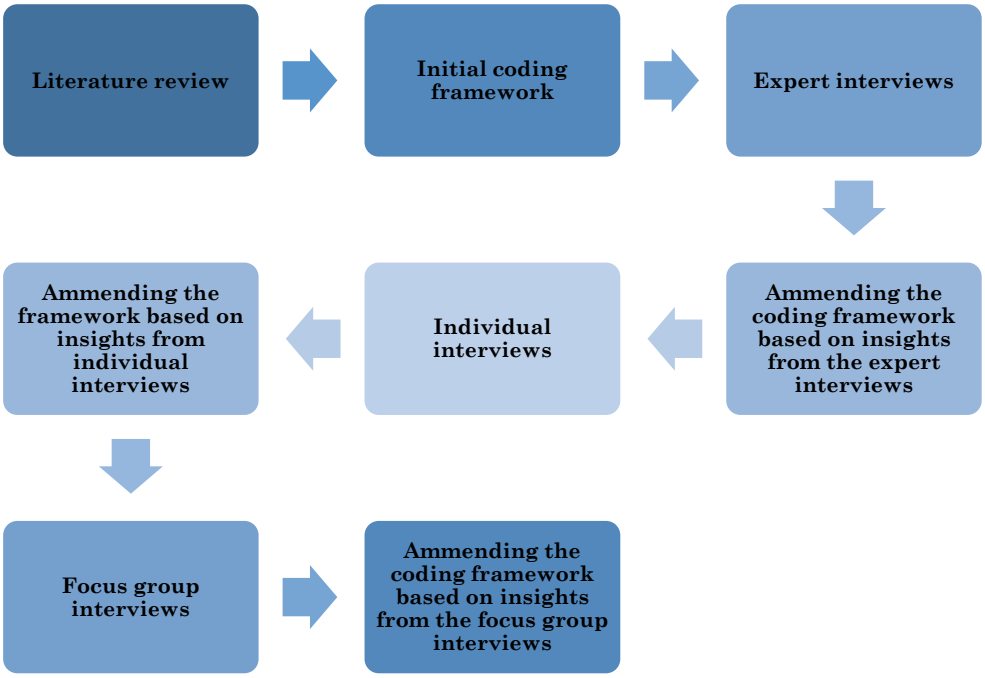

Figure 1: Illustration of the abductive approach used for the qualitative data analysis.

3.3. Research Strategy: Case Study of Imerco

To limit the scope, the present research uses the webshop of the Danish company Imerco as a point of departure, thereby grounding the findings in this reality (Saunders et al., 2009). Imerco was chosen due to convenience, since it was possible to edit their website design and obtain access to their Google Analytics account[12]. Imerco is a widely known Danish brand that generates a considerable part of its revenue online. It is the largest offline hardware chain in Denmark with 150 physical shops, and its online

[12] This was possible because Imerco is a client at the researcher's work place, the digital marketing agency Resolution Media.

revenue has grown exponentially since 2012[13]. It sells kitchen equipment, tableware, electric utility items and home decorations. Around 78.2% of Imerco's users are female and 21.8% are male, 53.35% are below 45 years old and 46.65% are above. 37.38% live in the capital region[14].

In the next section, theory about mood and judgment, and the quantitatively derived hypotheses that build on this theory, will be presented. This will be followed by theory on methods for improving mood, which drove the initial coding template for the qualitative analysis.

4. Theoretical Underpinnings

4.1. Our Reasoning Depends on How We Feel

"(...) the consciousness leaps into the magical world of emotion, plunges wholly into it by debasing itself. It becomes a different consciousness confronting a different world" (Sartre, 2004, p. 51)
"*Reason is, and ought only to be the slave of the passions, and can never pretend to any other office than to serve and obey them*" (Hume, 1896, p. 415)

Table 1: Quotes from philosophers about how our reasoning depends on how we feel.

People construe the world in different lights depending on how they feel (Oatley et al., 2006). According to K. Oatley, D. Keltner and J. Jenkins (2006), this "magic transformation" of the cognition by our affections has been depicted by artists and philosophers and assumed by laypeople for centuries (p. 259). In the above quotes by the philosophers Jean-Paul Sartre and David Hume (Table 1), it is described how our reasoning depends on our affections and how it is *not* possible to detach the two. During the enlightenment, most philosophers (with the exception of Hume) viewed this effect as a contamination of reason. Academically speaking, it has only been of interest during the last 40 years, but today researchers agree that there is no such thing as 'pure' cognition unaffected by affections (Clore et al., 2001; Forgas, 1995; Oatley et al., 2006).

Some researchers have dedicated their studies to examining how positive and negative mood states affect judgment (e.g. Forgas, 1995, 2009; Isen & Means, 1983; Park, Lennon, & Stoel, 2005; Schwarz & Clore,

[13] This data has been drawn from Imerco's Google Analytics account and is based on tracking from 2012 and up until today.

[14] This data has been drawn from Imerco's Google Analytics account and is based on tracking in the period 27-12-2014 to 27-12-2015.

1983; Schwarz, 2011). Other studies, collectively known as Appraisal studies, have focused on the effect of specific emotions, e.g. fear, sadness, gratitude and anger, on judgment (e.g. DeSteno, Li, Dickens, & Lerner, 2014; Lerner, Han, & Keltner, 2007; Lerner, Li, & Weber, 2013; Lerner, Small, & Loewenstein, 2004; Lerner & Keltner, 2000; Lerner & Tiedens, 2006). *Emotions* rise from implicit appraisals of situations. They are more intense and short lived, and they usually have a clear cause or referent (Forgas, 1995; Johnson & Stewart, 2010; Schwarz, 2011). *Moods* are of lower intensity and longer longevity. They are more generalised affective states that convey generic valence information and lack a clear cause or referent (Forgas, 1995; Johnson & Stewart, 2010; Schwarz, 2011). The terms *affect* and *feelings* are generally used to describe a generic category that accommodates both mood and emotion (ibid.). While mood researchers have focused on both the effects of positive and negative moods, most appraisal researchers have focused on negative emotions. Only few have studied positive emotions (e.g. Cavanaugh, 2009; DeSteno et al., 2014).

In Joseph Forgas' Affect Infusion Model (1995), he argues, based on a review of research from the 1980s and 1990s, that since emotions are more associated with conscious, motivated processing, whereas mood serves as unconscious inputs, emotions are less likely to influence the decision making process. This point is supported by various researchers who have found that mood only affects judgment if the subjects are not aware of the real source of their mood (Clore & Parrott, 1994; Schwarz & Clore, 1983; Schwarz, 2011; Siemer & Reisenzein, 1998). This phenomenon is known as the "misattribution" effect (Schwarz & Clore, 1983, p. 514) (cf. section 4.2.2.). Consequently, attention will only be paid to the effect of mood on judgment.

4.2. How Mood Affects Judgment

Researchers agree that mood affects people's cognitive evaluations and thereby biases the judgmental outcome - a phenomenon also known as "affect infusion" (Forgas, 1995, p. 39). This means that people evaluate objects and events more positively/negatively when in a positive/negative affective state (Forgas, 1995; Oatley et al., 2006). For example, in the classic study by N. Schwarz and G.L. Clore (1983), subjects where induced with a good/bad mood either by describing recent happy/sad events in their lives or by being interviewed on a sunny/rainy day. People in a good mood reported significantly higher life satisfaction than people in a bad mood.

People are influenced by their mood in all kinds of decisions in their daily lives, and the effect is not limited to any specific domain or to inconsequential judgment (Schwarz, 2011). For example, "public mood" due to weather conditions has been found to affect the stock market (Hirshleifer & Shumway,

2003). However, the effect is most likely to occur when people process new information, since they merely have to recall existing judgments when evaluating known information (Forgas, 1995).

Business researchers have found that good mood affects product evaluations and purchase intention positively. For instance, A. M. Isen, T. Shalker, M. Clark and L. Karp (1978) found that subjects in a positive mood[15] rated product performance and service significantly higher than subjects in a neutral mood. Similarly, G. J. Gorn, M. E. Goldberg and K. Basu (1993) found that people in a good mood evaluate stereo speakers significantly higher than people in a bad mood[16].

When asking subjects to imagine being highly involved in a shopping scenario (talking to a sales clerk), R. Swinyard (1993) found that people in a good mood[17] rated their general purchase intention (for any products in the store) higher than people in a bad mood.

The same conclusions were drawn by R. J. Donovan and J. R. Rossiter in two studies, one examining the effect of store atmosphere on purchase *intention* (Donovan & Rossiter, 1982) and one examining the effect of store atmosphere on purchase *behavior* (Donovan, Rossiter, Marcoolyn, & Nesdale, 1994). They found that the store atmosphere affected the customers' intended and actual amount of spending in the store, and that this effect was mediated by the level of pleasure experienced by the customer. Likewise, K. Spies, F. Hesse and K. Loesch (1997) tested the effect of mood on actual purchase behavior by comparing the amount of money spent on spontaneous shopping[18] by people visiting two different Ikea stores, one that had an unpleasant atmosphere[19], and one that had a pleasant atmosphere. The last group spent more money on spontaneous purchases and this effect was mediated by an improvement of their mood. Similarly, A. M. Isen and N. Geva (1987) argues that people in a positive mood are more willing to take small risks (Isen & Geva, 1987).

As previously stated, the number of studies on mood on judgment in an online setting is limited. J. Park, S. J. Lennon and L. Stoel (2005) tested the effect of product presentation on mood and purchase intention in an online apparel webshop. In a lab experiment, they presented 144 subjects with different variations

[15] Good mood was manipulated by giving the subject a gift. This method is described in detail in section 4.5.2.1.

[16] Mood was induced by listening to either mood-elating or mood-dejecting music.

[17] Good/bad mood was induced by varying the difficulty of an anagram test that the subjects were asked to perform. Section 4.5.1.2. explains this in detail.

[18] Spontaneous shopping is defined as purchases of things that one did not intend to buy before entering the store.

[19] The pleasantness of the store atmosphere was decided by lighting, brightness, colours and height of the room, which are considered important features for creating store atmosphere (Buckley 1987 as cited in in Spies, Hesse, & Loesch, 1997).

of a mock-up website presenting a pair of Khaki pants. In the variations, image size[20] and product movement[21] were manipulated. Although image size had no effect, putting the photos into motion raised the users' purchase intention for the Khaki pants and lowered their perceived risk of shopping on the website. This effect was mediated by an improvement in mood.

In the following, theories on how and why affect infusion takes place will be explained.

4.2.1. Affect Priming

The Affect Priming model was brought forth by Isen et al. (1978) and G. H. Bower (1981), and rests on the assumption that feelings (that is both emotions and moods, as explained in section 4.1.) *indirectly* affect the decision at hand by increasing the accessibility of affect-congruent information in memory. According to the theory, people consider both pros and cons based on memory, and positive feelings function as a cue to recall more material that is positive. Isen et al. suggest that this process can trigger a "cognitive loop" because mood triggers positive thoughts and positive thoughts in turn contribute to the maintenance of a good mood (p. 2).

Based on a number of experiments, Bower has put forward an Associative Network Theory, which suggests that memory is structured as an associative network in which feelings are central nodes that are linked to ideas and events of equal valence. Whenever instances of a feeling are experienced, they are recorded and tied to associations in memory, thereby connecting each event to a number of concepts. Whenever that feeling is stimulated in the future, those associations will be activated again (Bower, 1981). Thus, according to this theory, feelings can "prime the encoding, retrieval, and selective use of information" in cognitive processing (Forgas, 1995, p. 44).

4.2.2. Feelings as Information

The Feeling-as-Information model was brought forth by N. Schwarz (2011) and G.L. Clore, K. Gasper and E. Gavin (2001) among others. According to the Feelings-as-Information model[22], rather than computing a judgment based on recalled features of a target, people attend to their feelings as a heuristic

[20] Image size was manipulated by presenting one group with large photos and another group with small photos.

[21] Product movement was manipulated by presenting one group with front, side and back photos of the model and another group with one photo of a model rotating on a round plate.

[22] This model is also known as the Affect as Information model.

cue[23], which is a *direct* source of information. When we have to evaluate something, we simply ask ourselves: "How do I feel about it?" (Clore et al., 2001, p. 4). For example, if you are attracted to another person, the cause of the liking is how that person makes you feel – not the characteristics that you attribute to that person (e.g. friendly and kind) (Clore et al., 2001). This rests on an assumption that all feelings serve as direct feedback that guides judgment, decision-making, and information processing (ibid.).

People constantly stream their surroundings for affective cues in the form of feedback about progress toward minor sub goals[24] (ibid.). Usually the meaning of a cue is clear and using feelings as information therefore functions well, because the feelings are correctly attributed (Clore et al., 2001; Schwarz, 2011). However, people are more sensitive to their feelings than to the source of them, and they generally assume that their feelings are about that which is in the focus of attention. Consequently, feelings that are in the background, such as mood, which originally stem from one object or situation (e.g. the weather), are sometimes subconsciously attributed to another object or situation (Clore et al., 2001; Schwarz, 2011). In other words, people attribute their mood to whatever is in front of them, instead of analysing where it truly originates. Moods can thereby be "misattributed" to another source (Schwarz & Clore, 1983, p. 514).

Incidental feelings only inform judgment as long as the subject is not aware of the real source of the feeling. If the subject becomes aware of the source of their mood, it does not affect their judgment (Clore et al., 2001; Schwarz, 2011). For example, in G. L. Clore and W. Parott's study (1994), after being induced with a feeling of uncertainty[25], half of the subjects were made aware of the cause of their feelings. Then, they were asked to read and rate their understanding of a poem. Those who were not aware of the cause of their uncertainty, misattributed it to not understanding the poem[26], whereas those aware of the source did not let their feelings affect their judgment (Clore & Parrott, 1994). M. Siemer and R. Reisenzein (1998) also found that the effect of mood on evaluations of life satisfaction is less pronounced if the mood cause is made salient[27].

[23] Heuristic cues are easily noticed and understood cues that are linked to well-learned intuitive rules of thumb (Chaiken & Ledgerwood, 2007; Kahneman, 2002a). This is explained in detail in section 4.3.

[24] E.g. comprehending the instructions on a package or missing a stop light.

[25] This feeling of uncertainty (a negative feeling) was induced through hypnosis.

[26] They rated their understanding of the poem significantly lower than those who were aware of the cause of their uncertainty.

[27] Mood was manipulated by recall of positive/negative memories while listening to happy/sad music. The mood cause was made more salient by asking the subjects to fill out a questionnaire with mood scales.

4.2.3. Evaluation of the Two Theories

Whereas the Feelings-as-Information and Affect Priming models are regarded as contradictory, Forgas (1995) argued in his Affect Infusion Model that the two are actually complementary, since they apply to affect infusion under two different processing styles. He argued that whereas the use of feelings as information takes place under heuristic processing, affect priming takes place under systematic processing (these are explained in depth in section 4.3). The main principle behind the model is that the more systematic the processing style is, the more susceptible it is to affect infusion. This claim has, however, been rejected by other researchers who have found that people in a good mood use a more heuristic processing style than people in a bad mood, and that affect infusion is more pronounced when subjects use a more heuristic processing style (Schwarz, 2011). For example, Siemer and Reisenzein (1998) found that the effect of mood on judgment intensifies with reduced processing capacity[28], which means that the less systematic processing there is involved, the more pronounced the affect infusion is. In other words, a good mood lowers the level of systematic processing involved and thereby makes people more susceptible to affect infusion. This process is best explained by the Feelings-as-Information model, because it explains how people use their feelings as heuristic cues (Schwarz, 2011). According to Schwarz and Clore (1983), the misattribution effect also demonstrates the Feelings-as-Information model's superiority to the Affect Priming model: If the effect of mood were due to a higher availability of mood-congruent associations, then the effect would not be erased by making the subject aware of the source. According to the Feelings-as-Information model, affect infusion takes place namely *because* the feeling is misattributed. Thus, affect infusion is best explained by the Feelings-as-Information model.

4.2.4. First Two Quantitatively Derived Hypotheses

Based on the Feelings-as-Information model, and on the large amount of research suggesting a positive relationship between mood and evaluations, it can be assumed that Imerco's users' evaluation of the website and the products on the website will be affected by their mood, since they will attribute their feelings to the object of evaluation. The better their mood is, the higher their evaluation will be:

H_1: The better the users' mood is, the more positively they will evaluate the design of Imerco's website.

[28] Reduced processing capacity was induced by adding time pressure and competing task demands.

H_2: The better the users' mood is, the more positively they will evaluate a product on Imerco's website.

4.3. Mood and Processing Style

According to Dual Process theory, which has been brought forth by S. Chaiken and C. Ledgerwood (2007) among others, the heuristic and the systematic processing style represent two distinct systems of judgment. System 1 (heuristic) is a fast, automatic and spontaneous system based on intuitive associations, and system 2 (systematic) is a deliberate, in-depth and effortful step by-step mode (Chaiken & Ledgerwood, 2007; Schwartz, 2009). When people use a highly heuristic processing style, they do not deliberately think about the topic at hand, whereas when they use a highly systematic processing style, they attempt to thoroughly understand it (Chaiken & Ledgerwood, 2007; Kahneman, 2002).

J. T. Cacioppo and R. E. Petty (1984) suggested that people's motivation and processing ability determine how deeply they process arguments. This is determined by various conditions[29] – for example by time pressure and competing task demands, as suggested by Siemer and Reisenzein (1998). What makes processing style particularly interesting to the present research is the fact that mood is one of these conditions. It has generally been agreed that a positive mood leads to a more heuristic processing style, and that a negative mood leads to a more systematic processing style (Bless, Bohner, Schwarz, & Strack, 1990; Schwarz, 2011; Tiedens & Linton, 2001). As with the direct informational effect of feelings on judgment, the influence of feelings on processing style is also eliminated if the feeling is correctly attributed to its source (Schwarz, 2011).

People are most prone to use a heuristic processing style, since it is the least effortful option (Forgas, 1995; Schwarz, 2011). However, the more things feel wrong and do not go smoothly, the more attention people pay to detail in order to determine what went wrong (ibid.). On the other hand, good moods impairs attention to detail and processing precision, because it informs people that the situation is favorable and that little processing effort is needed (Schwarz, 2011).

[29] For example, it has been found that the level of systematic processing involved depends on how personally consequential the outcome is, whether there are any distracting tasks and how much prior knowledge the subject has about the issue at hand (Cacioppo & Petty, 1984).

4.3.1. The Third Quantitatively Derived Hypothesis

Based on the above-mentioned research, it is expected that if the mood of the users at Imerco's webshop is lifted, then they will use a less systematic (more heuristic) processing style. This leads to the next hypothesis:

H_3: The better the users' mood is, the less systematic their processing style will be.

4.4. Mood and Persuasion

The result of processing is used to guide subsequent attitudes, judgments, and behaviors (Chaiken & Ledgerwood, 2007). As mentioned, the finding that a good mood leads to a more heuristic processing style is important to the present research because it has been found that the effect of mood on judgment is more pronounced when a heuristic processing style is used (Schwarz, 2011; Siemer & Reisenzein, 1998). In addition, it is relevant that people who use a heuristic processing style have been found to be easier to persuade with weak arguments (Bless et al., 1990; Bless, Mackie, & Schwarz, 1992; Mackie & Worth, 1989). One reason for this is that people who use a heuristic processing style do not reflect on the quality of arguments when processing them. Furthermore, even if a message has been processed systematically, if people are using a heuristic processing style at the time of evaluation they still fail to retrieve the details of strong arguments. In that way, processing style affects judgment of information, both at the time of processing and at the time of evaluating (Bless et al., 1992). This is because people rely on heuristic cues when forming the judgment (Bless et al., 1992; Mackie & Worth, 1989). Heuristic cues are easily noticed and understood cues that are linked to well learned intuitive rules of thumb (Chaiken & Ledgerwood, 2007; Kahneman, 2002). These cues could for example be the communicator's credentials (e.g., expert or not), group membership, or - as suggested by the Feelings-as-Information model - positive feelings (Schwarz, 2011). Instead of going into critical depth with evaluating the arguments, people who use heuristic processing trust the intuitive response that comes from reliance on the heuristic cues in order to make quick and effortless decisions (Chaiken & Ledgerwood, 2007). In other words, when affect infusion takes place, people attend to their positive feelings as a heuristic cue (Schwarz, 2011).

4.4.1. The Fourth Quantitatively Derived Hypothesis

Based on the above-mentioned research, it is expected that if the users at Imerco's webshop are in a good mood and thereby use a more heuristic processing style, they will be easier to persuade to buy products

online. This is because they will not go into critical depth with evaluating the product attributes and prices. In addition, when asked to evaluate their purchase intention, they will use heuristic cues such as their positive mood, the brand's credibility and the sales labels for making the decision. This leads to the last quantitatively derived hypothesis:

H_4: The better the users' mood is, the higher their purchase intention will be.

Now, all expectations about how users' judgments are affected by their mood have been presented. These constitute the foundation of this research, since they justify the need to know how to affect the users' mood. The qualitative study seeks to answer this question. Before conducting the qualitative study, an initial literature review was conducted on the various methods that previous researchers have used for improving mood in other contexts. As the qualitative template analysis was initially based on concepts derived from this theory, it will be presented in the following section.

4.5. Mood Induction Methods

Mood researchers have used various methods for manipulating their subjects' mood. These are called MIPs (Mood Induction Procedures). Some of the procedures may also have an effect in the present context. In the following, findings regarding elements that have been found to affect mood in online contexts are presented firstly. Then, findings about offline MIPs that may be applicable online are described.

4.5.1. Elements that Have an Effect on Mood in Online Contexts

Even though research on the affective role of different website elements is limited, a few researchers have confirmed the effect of certain methods for affecting mood in an online context. These are presented in the following.

4.5.1.1. Facial Expressions

D. Cyr, M. Head, H. Larios, and B. Pan (2009) found that human images with facial features induce users to perceive a website as more appealing, having warmth and social presence. Even though they did not measure the effect on mood, these findings suggest that human images may also affect mood on a website (Cyr, Head, Larios, & Pan, 2009). This method has also been used in offline contexts. For example, Winkielman, Zanjonc and Schwarz (1997) found that photographs of individuals expressing anger or happiness deteriorates/improves mood (Winkielman & Zajonc & Norbert Schwarz, 1997). On a website,

this MIP can be applied by use of photos of people smiling, but it may also be applied by use of smileys[30] in the communication.

4.5.1.2. Photos and Cartoons

In an online survey experiment, A. S. Göritz (2007) tested six MIPs that have traditionally been used offline. The fact that the MIPs were performed on a screen makes the study similar to the present context, although it still differs from the natural context of a website. Göritz confirmed that pleasant photographs[31] and cartoons have positive effects on mood. Photos have also been used as MIPs in offline contexts. For example, Greenwald, Cook and Lang (1989) meassured affect with facial electromyography (EMG) and found that positively valenced photos[32] raised the activity in the muscles responsible for smiling (Greenwald, Cook, & Lang, 1989). Cartoons have only scarcely been used in an offline context (Göritz, 2007). On websites, they may be implemented in form of small cartoon company mascots. It should be mentioned that Göritz (2007) also calls for further exploration of possible online MIPs, since there may be other MIPs that are more suitable for the internet. The present research seeks to answer the call.

As previously mentioned, Park et al. (2005) found that making the clothing model in an apparel webshop spin around (product movement) can positively affect users' mood, which in turn leads to a lower perceived risk and a higher purchase intention. However, this is not relevant to the present research, as Imerco does not sell clothes.

4.5.2. Offline Mood Induction Methods

In an offline context, various MIPs have been used, some of which are applicable to the present online context[33]. These are described in the following.

---

[30] In this report, a smiley is defined as an icon or a few compounded characters that depict a facial expression.

[31] The pleasant photographs covered the themes; families, desirable women, comfortable home, happy children, fulfilled partnership, success at sports, unspoiled nature, cuddly animals, holidays, and children and animals.

[32] The positively valenced photos covered the themes; a happy baby, the opposite sex nude, a rabbit, a chocolate brownie and a dog.

[33] A number of MIPs are not applicable to the present context because they would interrupt the users' natural shopping flow. These are: 1. The Imagination MIP, where subjects induce their own mood by imagining situations from their lives that have previously evoked a happy/sad mood (Gerrards-Hesse et al., 1994). 2. The Velten Technique, where subjects are instructed to get in a certain mood while reading 60 self-referent statements intended to be elating/depressing (e.g. "I have too many bad things in my life" or "This is great – I really do feel good – I am elated about things") (Velten, 1968:475). 3. The film/story MIP, where subjects are presented with positive/negative narrative descriptive material and asked to get involved in the situation depicted (Gross & Levenson, 1995; Westermann et al., 1996). 4. The social interaction MIP, where subjects interact with a person that is behaving in a sad/happy way (Westermann et al., 1996).

4.5.2.1. Gifts

Offering subjects a gift was the method used in the already-mentioned study by Isen et al. (1978). Here subjects visiting a shopping center were offered a free product gift by another person only minutes before being approached by the researcher. This method could work in an online context as well, since a gift could be added to the users' shopping cart as a surprise loyalty bonus.

4.5.2.2. Feedback

Another popular MIP is providing the subject with positive/negative feedback after completing a task. In a study by A. M. Isen and T. E. Shalker (1882), the subjects were given a standardised test presented as a perceptual-motor problem-solving skill. Subjects in the positive/negative mood condition were told that their score was among the very best/lowest. In a similar fashion, Swinyard (1993) gave the subjects an anagram test accompanied by a text[34] describing how well others had performed. Mood was manipulated by varying the difficulty of the anagrams. This MIP can be integrated on a website by providing the users with feedback on completing goals on the website[35].

4.5.2.3. *Music*

Positive or negative mood-suggestive music has been used to induce a mood (Gorn et al., 1993; Westermann, Stahl, & Hesse, 1996). This element can be integrated on a website by playing music in the background when the website is used.

4.5.1.6. *The Most Effective Offline MIPs*

A. Gerrards-Hesse, K. Spies, and F. W. Hesse (1994) compared the mean effect size across nearly 250 mood studies from the previous ten years and found that the feedback and gift MIPs were more effective than the music and facial expression MIPs (the photo MIP was not included in their research). They also found that it is generally easier to deteriorate people's mood than to improve it, which may be due to the fact that baseline mood is generally biased in a positive direction.

To increase the effect size, some researchers have combined several MIPs (Gerrards-Hesse, Spies, & Hesse, 1994). For example, Siemer and Reisenzein (1998) asked their subjects to recall positive or

[34] The text read: "In previous tests of these words, the average upper-division business student has correctly solved 15 of the 25 anagrams within 5 minutes; 95 percent get at least 10 correct, and only the top 5 percent of students get 20 or more correct" (Swinyard, 1993, p. 274).

[35] E.g. the goal of finding products that allow them to save money or the goal of qualifying for free shipping.

negative memories while at the same time listening to either sad or happy music. This suggests that combining several website elements may also increase the effect size in an online context.

4.5.3. Initial Coding Framework for the Qualitative Analysis

To summarise, the MIPs that have been found to be effective and applicable to the online context of the present research are:

- Gifts
- Feedback
- Photos
- Music
- Facial Expression
- Cartoons

These MIPs constituted the codes in the initial template for the qualitative analysis. As mentioned, the purpose of the qualitative study was to explore how various website elements influence users' mood. In the next section, the methodology used for this study will be described.

5. Qualitative Methods

5.1. Expert Interviews

5.1.1. Methodology

The two expert interviews were conducted in the beginning of the study to gain knowledge from their practical experience (Kolb, 2008). In the following, the two experts[36] are presented.

5.1.2. Introduction of the Experts

5.1.2.1. Randall Rattan: User Experience Designer at Momondo

Randall Rattan is a User Experience Designer at the Danish company Momondo, a travel website and app that compares prices on flight tickets. According to the Head of Online Marketing, the company aims directly at affecting users' feelings positively through web design. The company has around 4-6 million users per month of which over half are returning users - a result that they attribute partly to their unique design that makes the users feel good (Kasper Hove, personal communication, March 5, 2015). By using

[36] The term 'expert' is used to refer to the two people who took part in the expert interviews.

ethnographic methods and usability testing, they seek to understand their users, which enables them to appeal to their feelings and thereby build a long-term relationship with them (Appx. 2). Randall works with improving the user experience of Momondo's website and app. He has worked with user experience design for almost 20 years at Momondo and other companies[37].

5.1.2.2. Michael Kjeldsen: Conversion Rate Optimisation Manager at Resolution Media

Michael Kjeldsen works as a Conversion Rate Optimisation Manager at the global digital marketing agency Resolution Media and is considered the Danish expert on web design and web analytics. He has extensive experience in the field, having worked with web design and analytics for almost 20 years. His primary job is to optimise the websites of Resolution Media's clients with the aim of increasing conversion rates[38]. He has extensive experience with usability testing and testing different variations of designs to identify the ones that drive the most conversions. His clients include a large number of global companies[39].

5.1.3. Instrumentation

A semi-structured interview guide was used as a guideline to make sure to stay within the main topics of interest (Appx. 1). The guide also allowed the experts to bring up new topics of relevance and enabled the interviewer to omit questions that had already been answered (Saunders et al., 2009). Since the experts were secondary sources (users being the primary sources to this study) the purpose was not to gain insights into the underlying meaning of their answers, why probing was only used for enhancing the validity by making sure that the information was understood correctly (Kolb, 2008; Saunders et al., 2009). The interviews were recorded to allow the interviewer to focus on asking questions, listening and showing interest. The interview with Randall was transcribed fully (Appx. 2), since all the information he provided was relevant. Michael often got off track during the interview (e.g. he started guessing what might affect the users instead of talking about his experiences) and therefore it was sufficient to replay the recording, take thorough notes, and write down quotes from the relevant parts of the interview (App. 3).

Interview bias was reduced by mostly asking open-ended questions that did not encourage the experts to answer in a particular way, but welcomed them to provide an extensive developmental answer (Saunders et al., 2009). By the end of the interview, the experts were given a list of various website elements on

[37] Sabre, Nokia, Microsoft and Navision.
[38] The percentage of visitors who perform a specific action such as buying a product or signing up for the newsletter.
[39] E.g. Bang & Olufsen and McDonalds.

which all the MIPs identified in the previous literature review were included. The experts were asked to identify and talk about the ones they knew and had not already brought up themselves (ibid.).

5.2. Individual Interviews

5.2.1. Methodology

Five individual interviews were conducted to gain a deep understanding of users' intuitive affective reactions and their conscious and subconscious associations with Imerco's current website design. The first part of the interview was a think-aloud user test designed to gain insights on the users' intuitive affective reactions and thoughts about the website. The second part was a card game used for mental model elicitation (explained in section 5.2.2.2.), which was used to gain deeper insights on how the users understand Imerco's site. Moreover, the card game served as a way of getting access to subconscious positive and negative associations and affective reactions related to various website elements. These are described below.

5.2.1.1. Think-aloud

In a think-aloud interview informants are asked to share their spontaneous thoughts while trying to solve a task (Buxton, 2007). In this case, the task was visiting a number of pages on Imerco's website. The purpose was to understand what the informants were thinking and thus gain insights into their spontaneous affective reactions and associations (ibid.). Think-aloud interviews are often used to understand users' problem solving strategies regarding a product, but they may also provide insights into feelings, which are unconsciously and uncontrollably activated when people verbalise their thoughts (Buxton, 2007; Hansen, 2005).

Since feelings are difficult to articulate, observations are valuable. This point was stressed by the expert Randall, who considers both observation and think-aloud interviews as important methods for gaining insights into the users' feelings:

"We see people's reactions and kind of like observe eye movement, facial expressions, body movement and things like that to kind of get the idea. I mean most times it's not difficult to look at a person to see how they react to something. But it's more difficult to ask them 'how do you feel about that?' and sometimes they may not want to share how they feel about it. So that part has always been kind of hard to measure, I think" (Appx. 2, p. 2).

Consequently, with the consent of the informants, the video cameras on their computers were turned on during the think-aloud interviews. When replaying the videos it was possible to observe the informants' bodily reactions (e.g. facial expressions) to different aspects of the website that they had not been able to articulate in the interview. Sometimes the informants moved their heads closer to the screen because they had trouble reading the text. This was an important observation related to the 'fluency' element discovered in the qualitative data (cf. section 6.1.3.).

5.2.1.2. Mental Model Elicitation

Mental models are people's internal iconic representations of the world, which form the basis of their reasoning. They are created little by little based on the meanings people ascribe to things through experience and knowledge (Johnson-Laird, 2010; Jones, Ross, Lynam, Perez, & Leitch, 2011). Mental model elicitation was used to gain an insight into the positive and negative meanings users ascribe to various elements on Imerco's site. It provided insights on their understanding of the website and the meanings they ascribe to various parts. It involved asking them to describe the website and sort different website elements based on categories *they* find meaningful, asking them to identify the elements *they* find most important and probing for explanations as to *why* that is. This will be explained in detail in the next section (cf. section 5.2.2.2.).

5.2.2. Instrumentation

Five informants were recruited by use of convenience sampling. The interviews took place in the informants' own homes to make them feel comfortable and make the use of Imerco's website appear like a natural everyday task. A semi-structured interview guide (Appx. 4) was used to stay within the main topics but also allowing flexibility. All interviews were recorded and replayed afterwards for thorough notetaking (Appx. 5).

5.2.2.1. Think-aloud

Informants were asked to perform a number of tasks on Imerco's website while imagining they were looking to buy new kitchen equipment. The tasks formed a realistic shopping scenario[40] to make the experience appear natural. The interviewer guided the informant through the tasks by reading each step

[40] The shopping scenario: Firstly, the informants were asked to go to Imerco's website and skim through the front page. Then, they were asked to find an overview of plates, skim through the first four pages of the product selection, select one they liked and add it to the shopping cart. Subsequently, they were asked to do the same with kitchen bowls and then the same with either gifts or blenders. Lastly, they were asked to go to the shopping cart and follow the payment steps while pretending that they were buying the products.

of the interview guide[41] (Appx. 4). The informants were asked to say aloud everything they thought about while performing the tasks (even those unrelated to what they saw) and to share how performing the tasks made them feel. They were told that there were no correct or incorrect answers. Probing was used continuously to make sure that the informants shared all their thoughts[42] and to gain deeper insights into why they thought or felt as they did (Appx. 4). Afterwards, the informants were asked to evaluate their experience of the website to gain insights into what they in hindsight considered to be good and bad experiences.

5.2.2.2. Mental Model Elicitation

The mental model elicitation method 3CM was used. This is a variation of the free card sorting method, which involves asking informants to sort and categorise items related to a particular subject and probing as to why those categorisations are made (Kearney & Kaplan, 1997).

Firstly, informants were asked to explain the content of the website as they would explain it to someone who had never seen it before. The explanations provided insights on the informants' impressions of the website and made it possible to understand how they structured those impressions (ibid.). Normally the card game sorting method fails to identify important aspects of informants' mental models, because the interviewer chooses the cards. Consequently, cards were chosen by the informant by asking them to list all the elements of the website that they remembered[43]. Informants were then asked to sort and categorise the cards in whatever way they thought to be suitable (ibid.). Another limitation of the regular card sorting method is its failure to identify the concepts informants conceive to be the most relevant. Therefore, in a second round of sorting, the informants were asked to arrange the cards according to how important those elements were to their shopping experience (ibid.). To make this easier, they were asked to write a number from one to five[44] on each card that represented how important they considered the element to be. In order to make sure that the informants had not forgotten any important elements, at the end of the interview they were given a pile of pictures with the elements that they had seen in the shopping scenario (buttons, photos, texts, etc.). The informants were asked to go through the pictures quickly and identify if there were any of the elements that they found to be very important but had forgotten to mention earlier.

[41] E.g. "Find an overview of all the plates that they sell and have a look at the page" (Appx. 4).

[42] E.g. "What are you thinking about now?".

[43] It was up to them to decide what they considered to be 'elements'.

[44] 1 being 'not important at all' and 5 being 'very important'.

5.3. Focus Groups

5.3.1. Methodology

A focus group interview is a formal way to make a group discuss the topic of interest (David & Sutton, 2007). It gives access to people's own interpretations of their affective behavior, but does not reveal their unrecognised affective reactions or capture the affective experiences of carrying out activities 'in action'. Thus, the purpose of the two focus group interviews was to confirm/reject the website elements already identified in the previous interviews and further explore the informants' own interpretations of them by providing quick access to a large variety of feedback from many users.

5.2.3. Instrumentation

Informants were recruited by use of convenience sampling. Five people had signed up for each interview, but in both cases only four appeared, which caused a skewed gender distribution in both groups[45]. The four informants in the first group had participated in the individual interviews as well[46]. The small size of the groups was an advantage because as a single interviewer it is challenging to both manage the process of the interview and keep track of all the information (Saunders et al., 2009). To manage the challenge further, the interview was recorded. During the interview, notes about nonverbal communication were taken. The records were transcribed subsequently to allow for precise coding of the data (Appx. 7).

To facilitate an open and honest dialogue about affective experiences (sensitive information) an effort was made to make the informants feel comfortable[47] and let them know what was expected of them (Saunders et al., 2009). Both interviews were initiated by two opening questions[48] that allowed the informants to take the discussion in a direction they felt comfortable with in order to get them started with expressing their thoughts and feelings. Afterwards, the informants were presented with a number of Power Point slides that presented website elements (Appx. 6). As the interview went along, the informants were asked to discuss their immediate affective reactions to the elements and not over-analyse their response. The slides were slightly modified for the second focus group in order to integrate the findings from the

[45] A male had to withdraw from the first one and a female had to withdraw from the second. Therefore, there were three females and one male in the first focus group, and three males and one female in the second.

[46] The informant who participated in the fifth individual interview (Appx. 5.5) was the only informant from the individual interviews who did not participate in the first focus group as well. He withdrew due to stress.

[47] The interviews took place in a large meeting room that was decorated to create a nice atmosphere. The informants sat around a small table with five seats in front of a projector. A wealth of refreshments was arranged on the table. Further, the informants were told that there were no correct or incorrect answers.

[48] "What makes you in a good mood online?" and "What makes you in a bad mood online?".

first group. The elements that the first group had clearly rejected were omitted, which allowed extra time for exploring the informants' reactions to different variations of the elements that had been confirmed.

A semi-structured interview guide was followed[49] (Appx. 6); the questions were open-ended and designed to encourage the informants to discuss their affective experiences in an open dialogue (David & Sutton, 2007). Probing was used to gain a deep understanding of the informants' answers, which enhanced the validity of the data (Saunders et al., 2009). The small size of the groups allowed the informants the time and space to express their thoughts[50].

5.4. Limitations of the Qualitative Study

5.4.1. Reliability Issues

In accordance with the interpretivist approach to the qualitative study, the findings are believed to be context dependent and representing the reality at the time of data collection (Saunders et al., 2009). Furthermore, researchers would not interpret the data in the exact same way, because they are biased by their own subjectivity (ibid.). Thus, an exact replication is not possible.

The bias of the researcher was reduced in the interviews by reacting as neutrally as possible to what the informants said to avoid affecting their subsequent answers. Yet being only one researcher on the project and not having the resources for an inter-rater reliability check may have increased the researcher bias. There is also a risk that the informants in the focus groups responded based on what they thought was the correct or socially accepted answer (ibid.). In the first group, most informants knew each other[51], which was an advantage because they appeared to be comfortable with sharing their thoughts and feelings. In the second group, most informants did not know each other[52]. This created a different dynamic, since the informants appeared to agree more with each other than the first group. The fact that the informants may have answered differently if they had been asked individually serves as an important consideration that may challenge the reliability of the data.

[49] The interview guide was included in the Power Point slides as notes that only the interviewer could see.

[50] Some were more eager to talk than others were, but due to the size of the groups, it was easy to probe everyone to share their thoughts, e.g. by asking, "Do you agree?" or just by sending an inquiring gaze and a smile.

[51] One informant had only met two of the other informants before, while all other informants knew each other.

[52] Only two of the informants knew each other.

5.4.2. Sampling Limitations

The lack of resources prevented use of purposive sampling. Consequently, the informants were selected haphazardly among friends and acquaintances. This made the sample prone to bias and influences beyond the researcher's control. Whereas Imerco's users are a rather heterogeneous group based on gender, geography and age (cf. section 3.3.), the informants who participated belong to a rather homogeneous group, since they are all 25-30 years old, have a university degree and live in Copenhagen. The sample is thereby not representative of the population, and subsequent generalisations will be flawed.

In the next section, the findings from the qualitative study will be presented.

6. Qualitative Findings

The results suggest that users are positively affected by certain types of photos and background colours, fluency of information, gifts and clear service and product advantages. These are henceforth called 'mood cues' because they influence mood positively. The empirical findings regarding the mood cues will be explained in this section in order to outline how and why these elements can be used to affect users' mood. The findings about elements that were included in the initial conceptual framework, but not empirically found to influence mood positively (cartoons, feedback and music), are presented in the appendix (Appx. 13).

6.1. Mood Cues

6.1.1. Photos

Photos[53] are the website elements that have the most obvious and strong influence on mood. However, this is the case for both positive and negative moods, and since people have various unpredictable associations with photos, choosing one that will influence mood positively is a challenge. The experts emphasised that photos are the quickest and most effective elements for affecting users' mood (Appx. 2 & 3), but also stressed that the valence of the users' reactions is difficult to control and predict (Appx. 2 & 3). As Michael phrases it, photos "yell at people"[54] because they catch the attention so easily (Appx. 3,

[53] This definition includes both background photos and front ground (regular) photos, since the informants did not differ much on their reactions to the two.

[54] The quotes from the expert Michael and from all the informants have been translated from Danish into English. Only the quotes from the expert Randall have not been translated, as they are already in English.

p. 2). This can be an advantage, if the users react positively to them, but unfortunate if the photo unintendedly provoke negative feelings.

All photos presented to the focus groups depicted happy or pleasant situations[55] (Appx. 6). The informants had strong affective reactions to the photos, but these were both positive and negative. Still, it was possible to identify some general patterns regarding the informants' affective reactions to different characteristics of photos. These are presented in the following subsections.

6.1.1.1. Photo Characteristics that Affect Mood Positively

In this section, insights regarding characteristics of photos that generally affect users' mood positively will be addressed.

6.1.1.1.1. Tranquility

"What I like about the photo with the coffee cup is the calm background", one of the informants said (Appx. 7.2, p. 13). The informants often commented positively on photos that they found "calm and pleasant to look at", "less distracting" or "subtle" (Appx. 7.2, p. 9). What generally characterised these photos was that they were toned down and used muted colours. The photos did not cause strong affective reactions, but relaxed feelings of pleasantness. In general, the simpler the photo the more pleasant the informants found it. For example, when presented with a background photo of trees in fall colours, the informants found it annoying and distracting. "It is too patterned, not calm", one of the informants said (Appx. 7.2, p. 5). When presented with a more simple photo depicting a sunset sky in muted and blurry light blue and orange colours, they liked it more because "it is more calm and pleasant to look at" (Appx. 7.2, p. 6). The expert Randall also mentioned that a website design needs to be "subtle and pleasing" (Appx. 2, p. 8). It needs to make the user feel "calm and comfortable" (Appx. 2, p. 2). Tranquility was especially important to the informants when it came to using background photos, because these are larger and therefore more noticeable (Appx. 7.1).

6.1.1.1.2. Aesthetics

When arguing why they liked or disliked photos, the informants sometimes referred to aesthetics, e.g. an informant disliked a photo because he did not find it neat: "It is annoying with the wrinkled tablecloth (…) and the bread does not look good on the chopping board, it does not fit..." (Appx. 7.2, p. 19). The

[55] E.g. a photo of a woman smiling while enjoying a cup of coffee, children laughing while playing, a cosy living room with candle lights and a sunset.

experts also emphasised that people are negatively affected by things they find ugly. According to Randall, Momondo has succeeded in creating a "nice and pleasant environment" by using colours and images that people generally find beautiful (Appx. 2, p. 2).

6.1.1.1.3. Unstaged

Informants reacted more positively to photos that they thought looked natural and non-commercial, and they could quickly detect if a stock photo had been used. "Who puts a lot of coffee beans into a coffee cup?", an informant said when seeing a photo of a coffee cup with beans on the page displaying espresso machines: "It is too staged and my mood gets worse just from looking at it" (Appx. 7.2, p. 9). The expert Michael also mentioned this issue: If photos do not look natural, they do not appear trustworthy (Appx. 3).

The informants were critical about this issue and even identified some photos that were *not* stock photos as stock photos. This was especially true for photos that were not product or company specific. Photos that did not have a clear purpose made the informants suspicious: "I would feel that they are trying to hide high prices by adding some nice photos and try to appeal to the users' emotions", an informant said (Appx. 7.2, p. 12). In other words, if it was obvious that the purpose of the photo was to influence their affections it actually had the opposite effect.

6.1.1.1.4. Accompanied by Text

When skimming though a page on the website, the informants in the think-aloud interviews generally took more notice of the photos that were accompanied by a text/headline than photos standing alone. This confirms the experts' experiences, since they both mentioned that a combination of text and photos is the strongest way of affecting mood (Appx. 2 & 3).

6.1.1.1.5. Displaying the Products in a Nice, Familiar and Decorative 'Dream' Setting

In general, the informants reacted negatively to photos depicting settings that they could not relate to, e.g. an informant was negatively affected by a photo of a sunset: "(…) this one looks like it was shot in Japan or something – far away from anything that has anything to do with us" (Appx. 7.2, p. 6).

In addition, the informants had positive affective reactions to photos depicting products in a nice decorative setting. In general, the informants liked the idea of seeing the product in a cool interior, e.g. a blender in a cool kitchen, a candleholder in a nice living room or a grill on a trendy terrace, because it made them feel they were buying something luxurious. The informants did not just want to buy the

product, but they also wanted to buy the dream of the product as part of a cool and luxurious home. "You need to feel that you don't just buy the product but also get a feeling or… (…) a dream that comes with it!", an informant said (Appx. 7.1, p. 10-11). The informants liked the idea of "stepping into a room on the website" (Appx. 7.2, p. 7) – an online room that looks the way that you would like your home to look[56]. "There is this strange and subtle satisfaction in seeing the product standing there on a cool table (…)", an informant said (Appx. 7.2, p. 7). Some also said that they liked imagining themselves in that setting, e.g. one informant said he would prefer a homely photo that made him imagine himself in his kitchen "having a good time, standing there cooking - and there are candle lights…" (Appx. 7.2, p. 19). This informant referred to the Danish concept of 'hygge' (best translated with the word cosiness), which describes a warm ambiance of having a good time with people you like. When commenting positively on photos, informants often used this word, especially when commenting on photos that depict homely settings or people socialising and having a good time. This suggests that photos that create this atmosphere have a positive influence on the mood.

The experts also found it important to choose motives that people recognise and have good associations with. For example, Randall said that when people see a candle light, they think of 'hygge' and that feels good to them (Appx. 2). According to him, photos should encourage people to use their imagination: "I think it's about triggering their imagination a bit in the right direction and letting them fill in the blanks" (Appx. 2, p. 9).

This finding relates to the previous finding by Göritz (2007) (cf. section 4.5.2.) that pleasant photographs affect mood positively. In addition, it is interesting that one of the photographs used in Göritz' experiment was also a photo of a comfortably looking home.

6.1.1.2. First Qualitatively Derived Proposition

The above five points lead to the first qualitatively derived proposition:

P_1: Users who are exposed to calm, neat, and unstaged photos accompanied by text and displaying the products in a nice, familiar, and decorative 'dream' setting will be in a more positive mood than users who are not exposed to these photos.

[56] They compared this feeling to walking through a furniture store where each room is decorated as a real room in a home.

The above section addressed characteristics of photos that generally affect users' mood positively. However, there are also some characteristics that the informants had mixed affective reactions to, which were therefore not included in the proposition. These will be shortly presented in the following.

6.1.1.3. Photo Characteristics that Elicit Mixed Reactions

6.1.1.3.1. People in Photos

Both of the experts stressed that users usually have stronger positive affective reactions to photos of people (Appx. 2 & 3). Randall thinks users like seeing pictures of people who they would like to be with: "You know we all want to be with people who look like they are having fun, smiling or laughing of something, or just enjoying themselves", he said (Appx. 2, p. 9). Both experts also believe that it is an advantage if the people in those photos smile because it makes them look friendly (Appx. 3 & 2), which confirms the facial expression MIP identified in the literature.

On the other hand, the informants had mixed reactions to photos with people in them.

The informants in the first group expressed strong positive affective reactions to photos of people because it helped them imagine making use of the products in a happy life situation. E.g. an informant said: "If I have to buy a new winter jacket and I see a photo of some people dressed in the fall outfit and being out in the woods looking like they think that life is wonderful (…) I would be affected by that (…)" (Appx. 7.1, p. 1-2). These photos had a more profound influence on their mood if they could relate to the situation depicted in the photo; e.g., a photo of three friends standing on a beach at sunrise moved two informants because it made them recall similar personal life situations. One of them said: "It is because I recognise myself in that feeling of summer (…). It just makes you happy (…)" (Appx. 7.1, p. 6). When presented with another photo of a family with a child, they were not affectively influenced by it because they could not relate to the situation (Appx. 7.1).

People in the second focus group generally disliked photos of people because they found them to be too commercial/staged. They did not believe the depicted life situation to be natural: "It can seem a little bit too fake when there are people in the photos", an informant said (Appx. 7.2, p. 10). Instead of letting themselves be carried away by the story told in the photo, they resisted being fooled by what they considered an obvious commercial manipulation. On the other hand, when presented with photos that looked less staged because they were not as professionally set up and retouched, some informants found the photos to be too personal: "It is a little bit like looking into someone's family album", an informant said (Appx. 7.2, p. 20).

The contradictory findings in the two groups suggest that the gender differences in the two groups may have affected the results. A possible explanation may be that women are more easily carried away by photos depicting affective life situations and do not detect the commercial purpose to the same extent as men. An alternative explanation is that the differences are not necessarily due to gender differences, but caused by reliability issues in the data from the second interview due to the social code[57].

6.1.1.3.2. Smileys

Another way to integrate the facial expression MIP on a website could be to use smileys. Both experts expressed that smileys could possibly affect people's mood positively, even though they had no data to support this claim. Momondo is actually using three variations of smileys[58] at their website to indicate the value for money on flight tickets, but this design choice was more a response to current trends than to user insights (Appx. 2). Nevertheless, the informants reacted negatively when seeing smileys on a website, because they found them to be unprofessional (Appx. 7.2 & 7.2).

6.1.2. Background Colour

According to the experts, the combination of colours and light on a website affect users' mood (4-5). This applies especially to the background colour on a website. However, neither of the two experts could provide an exact recipe for how various colours affect mood, as it depends on the context (Appx. 2 & 3). The 'right' colours for a specific website depend on its user. Not knowing the users, using colours can go terribly wrong[59] (Appx. 2).

The informants were more pleased with a background colour that created a contrast between background and foreground (Appx. 7.1, 7.1, 5.1, 5.2 & 5.5). When the informants tested Imerco's website, the majority noticed the white background colour (Appx. 5.1, 5.2 & 5.5). "The site is extremely white", an informant said (Appx. 5.1, p. 2). The focus groups also liked the website better with a background photo than with a white background (Appx. 7.2 & 7.1). Some did not like the white background because it made the other elements on the website merge into the background, while others said they found the colour boring (Appx. 7.2 & 7.1). On the other hand, the informants liked that the white background was clean, neutral and did

[57] The informants in the second focus group sometimes got carried away with being critical, possibly because they replicated and built on to each other's answers.

[58] Sad, indifferent and happy.

[59] Randall explained how he used to work for Microsoft who used many colours and icons in the web design because it was trendy back then. Later, the company learned that this actually had a negative impact, because the users found it childish (Appx. 2, p. 4).

not attract too much attention (Appx. 5.1, 5.2, 5.5, 7.1 & 7.2). In general, the lighter the colour[60] (but not white), the more calm and pleasant the informants found it (Appx. 7.1 & 7.2). Whether or not they personally liked the colour played a role too, e.g. two informants were drawn by a petroleum blue background colour. Interestingly, one of them was wearing a watch in the exact same colour (Appx. 7.2).

6.1.2.1. Second Qualitatively Derived Proposition

The above points lead to the second qualitatively derived proposition:

P_2: Users who are exposed to a website that uses a calm and light background colour that creates a contrast between background and foreground will be in a more positive mood than users who are exposed to a website that uses a clean white background colour that does not create a contrast between background and foreground.

6.1.3. Fluency

In general, the informants agreed that being able to skim through a website quickly and easily made them feel good, while they got annoyed if it was difficult or took too long to process information (Appx. 5 & 7). This cue is named 'fluency', because this term defines how easy it is to process information (Schwarz, Sanna, Skurnik, & Yoon, 2007). The experts also regarded fluency as an important factor of affecting the users' mood, since users feel better if they can quickly browse through all information and fully understand it without any hindrances (Appx. 2 & 3). Both informants and experts mentioned a number of factors that influence mood because they either disturb or enhance the processing of a website. These are explained in the following.

6.1.3.1. Not Too Much Information at Once

The informants disliked the existing product and category descriptions on Imerco's website, because they found most of the information to be unnecessary and annoying. A product description of drinking glasses said that the glasses are "perfect for a hot day in the garden", a description the informants said was "nonsense"[61] (Appx. 7.1, p. 23). The informants emphasised that they liked websites better when the information provided was "short and precise". Only factual information about the product and brand

[60] For example, they liked cream and light grey.

[61] In Danish, the users used the words 'pladder' and 'sludder', which have been translated as 'nonsense'. It means that the content of the text is trivial and unimportant filling material.

should be included (Appx. 7.1, p. 23 & 7.2, p. 23). The informants preferred information presented in bullet form, because this made it easier to skim through the text and made the presentation look more orderly (ibid.). Some suggested it would be better if the information was sorted in chunks where only the headlines were visible and the body texts were hidden and expandable[62] (Appx. 7.1). The amount of non-textual information, e.g. photos and buttons, also affected the processing ease. In general, the less information there was on the page, the more accessible, calm, and pleasant the informants found it to be (Appx. 5 & 7). The informants who tested Imerco's website found it messy and chaotic because there were many unnecessary and distracting elements that disturbed their processing (Appx. 5.1, 5.2 & 5.3). The elements that they found unnecessary and disturbing are listed in Table 2.

Elements on Imerco's website that the informants found unnecessary and disturbing

- Irrelevant offers displayed on the front page (irrelevant because they were not personalised and none of the informants found them useful) (Appx. 5.1, 5.3 & 5.5)
- Social share buttons at the product presentations (Appx. 5.3 & 5.5)
- 'Add to wish list' button at the product presentations (Appx. 5.3 & 5.5)
- 'Add a discount code' field (Appx. 5.2 & 5.5)
- EAN number on all products (Appx. 5.2 & 5.5)
- Ads for products they were not interested in (Appx. 5.1-5.3 & 5.5)
- Long titles of products that include too much information (Appx. 5.3)
- Order number and store number on the checkout page (Appx. 5.3 & 5.5)
- The information "status: in store" (since the informants expected all products to be in store if they were displayed) (Appx. 5.1, 5.2 & 5.5)
- Sign up element for Imerco club membership – the majority of the informants were not interested in this service, because they were first-time users (Appx. 5.1-5.3 & 5.5), and yet they were presented with the element several times, which one informant indicated was annoying (Appx. 5.5).

Table 2: List of elements on Imerco's website that the informants found unnecessary and disturbing

[62] This is called an "Accordion" (JQuery, 2015). By using this, the text would not take up space and the users could click on the headline and fold out the text only if they wanted to read that information.

This confirms what the experts said: That the website needs to be simple and only include the elements that are necessary for the users to make their decisions, because many users become annoyed if there is too much activity on a website (Appx. 2 & 3). "I think that a lot of sites they are counterproductive in that manner cause they have so many flashes and ads going on (…) so it's all about giving them the comfort and not distracting them", Randall said (Appx. 2, p. 7). Michael recently confirmed this in an online test where he removed many elements from a client's website in order to make the website look more simple, calm and clear. This change drastically increased the conversion rate (Appx. 3).

6.1.3.2. Clarity of Information

One of the experts emphasised the influence of size, colour and font of text on processing ease: To influence mood positively, the font has to be familiar and the words have to be clear against the background and large enough to be read easily (Appx. 3). That the informants reacted negatively to the small text size and small photos of the products by moving closer to the screen and squinting when they had to read text or evaluate photos (Appx. 5.1-5) confirmed this. In addition, they said that the small size of the photos annoyed them, because they expected clear, high quality photos in webshops (Appx. 5 & 7). Some also commented on the lack of contrast between some of the product photos and the background[63]. The photos have the same white background colour as the entire page and some products are even white themselves (e.g. plates) (Appx. 5.1 & 7.1).

6.1.3.4. Third Qualitatively Derived Proposition

The above-mentioned findings about fluency lead to the third proposition:

P_3: Users who are exposed to information that is easy to process because it is clearly presented (due to large size and high quality of text and photos and due to clear font and contrasts) and because only small amounts of information is presented at once, will be in a more positive mood than users who are exposed to information that is difficult to process.

6.1.4. Gift

Interestingly, the informants in the first focus group brought up the gift topic themselves when they were discussing what puts them in a good mood online. They discussed how they love getting surprisingly good

[63] This insight also relates to the background colour cue. However, the background colour cue has to do with colour, whereas this specific insight has to do with contrast (which relates to fluency).

treatment from companies. The informants mentioned receiving their order in a nice gift wrapping or discovering that the company has added an extra item as a surprise gift: "(…) then it is a bit like a present instead of just a box with a thing in", an informant said (Appx. 7.1, p. 4). Another informant said that each time she receives an order of razors, she knows there will also be a caramel in the package: "And it is such a small thing, but I know that when that small package arrives, there will be a caramel. And then I actually get a little bit happy", she said (ibid.).

All informants agreed that it would also make them happy if a gift was unexpectedly added to their online shopping cart: "I would be completely sold (…) I would be like 'oh, that is so nice of them…'", an informant said (Appx. 7.1, p. 18). The idea of receiving a gift made the online experience appear more personal to them. They compared it to the sensual experience of shopping offline and getting special treatment by the sales clerks – a sensual experience that is normally not possible to gain online. "I normally like shopping offline and the whole shopping experience that I enjoy I do not get that online, so online you need to get another kind of good experience", an informant said (Appx. 7.1, p. 24). "The gifts does the job. It brings a sensual experience home to you", another informant said (ibid.). They agreed that they would prefer if the gift were something that everyone needs – e.g. a vegetable knife (Appx. 7.1). They also stressed that if they received a gift, they would expect to receive one again next time they did their shopping on the same website unless it was made clear that the gift was a one-time thing (ibid.). A few informants expressed that receiving a surprise gift online might make them suspect the company of trying to make them buy more (Appx. 7.2). This suggests that it should be made clear that the gift is free and does not impose extra expenses on the user.

6.1.4.1. Fourth Qualitatively Derived Proposition

The above-mentioned findings about fluency lead to the fourth qualitatively derived proposition:

P_4: Users who are offered a gift will be in a better mood than users who are not offered a gift.

6.1.5. Advantages

According to the experts, service and product advantages make a website appear as more trustworthy (Appx. 2, p. 8 & 3, p. 3). Imerco communicates a number of service and product advantages on their site, which are listed in Table 3.

Service and product advantages communicated on Imerco's website

- Sales labels (communicating savings)
- Free shipping when you buy for more than 300 kr.
- 30-day return policy
- The option of returning the product in an offline shop
- Digital receipts
- Positive customer ratings

Table 3: List of service and product advantages communicated on Imerco's website.

The sales labels were easily noticed by informants in the think-aloud interviews and they generally reacted positively on the information when it was in relation to products they found interesting (Appx. 5.1-5). As mentioned in the section about fluency, some of the informants were annoyed by the prominence of the sales marks in relation to products they did not find interesting (cf. 6.1.3.1.). When asked about what puts them in a good mood online, the majority of the informants in the focus groups also mentioned good offers (Appx. 7.1, p. 2 & 7.2). They also reacted positively on the remaining brand advantages when they were asked about them, but they did not notice any of them by themselves when visiting the site. This suggests that they would be positively affected if the service and product advantages were made more noticeable.

6.1.5.1. Fifth Qualitatively Derived Proposition

The above-mentioned findings lead to the last qualitatively derived proposition:

P_5: Users on Imerco's website exposed to a product page with clear information about service and product advantages will be in a better mood than users on Imerco's website who are exposed to a product page without clear information about service and product advantages.

6.2. Individual and Contextual Differences

In general, the findings suggest that there are complex individual differences between users, because some respond differently to the same website elements (Appx. 7. & 5). In addition, it was found that the effect of various website elements depend on the context and the company's brand. For example, both informants and experts suggested that using funny website elements may have a positive effect if the brand has a fun and relaxed image, but not if it has a serious image (Appx. 2, 3, 7.1 & 13). This was the

reason why the users reacted negatively to the idea of including cartoons on Imerco's website: "Imerco should not be fun" (Appx. 7.1, p. 20).

6.3. A Mood-optimised Design

In spite of the above-mentioned individual differences, it was possible to identify some elements that influenced all of the informants positively. Each of the five propositions that emerged from the qualitative findings are based on a separate mood cue that can be expected to affect mood positively. These are summarised in Table 4.

Mood cues

- **Photos:** Calm, neat, and unstaged photos accompanied by text and displaying the products in a nice familiar and decorative 'dream setting'.
- **Background:** Calm and light background colour
- **Fluency:** Information that is easy to read
- **Gift:** Offering the user a gift
- **Advantages:** Clear service and product advantages

Table 4: Mood cues identified in the qualitative data.

6.3.1. The Qualitatively Derived Hypothesis

As previously mentioned, the goal of the present research was to integrate the qualitative insights into a 'mood-optimised' mock-up design and test it in an experiment to discover whether it raises purchase intention, product evaluation, website design evaluation and processing style. Due to resource constraints, it was not possible to test and confirm the effect of each of the mood cues individually, which meant that all five cues had to be tested together. The five qualitatively derived propositions were therefore merged into one qualitatively derived hypothesis:

H_5: Users who are exposed to a design that contains the five mood cues (the mood-optimised design) will be in a better mood than users who are exposed to a design that does not contain the five mood cues (the original design).

The five hypotheses that have so far been stated predict relationships between mood and design evaluation, product evaluation, processing style and purchase intention (quantitatively derived) and between design and mood (qualitatively derived). The final four hypotheses, thus, posit a relationship between design and design evaluation, product evaluation, processing style and purchase intention that is mediated by mood:

H_6: Users who are exposed to a design that contains the five mood cues (the mood-optimised design) will evaluate the design of Imerco's website more positively than users who are exposed to a design that does not contain the five mood cues (the original design).

H_7: Users who are exposed to a design that contains the five mood cues (the mood-optimised design) will evaluate a product on Imerco's website more positively than users who are exposed to a design that does not contain the five mood cues (the original design).

H_8: Users who are exposed to a design that contains the five mood cues (the mood-optimised design) will use a less systematic their processing style than users who are exposed to a design that does not contain the five mood cues (the original design).

H_9: Users who are exposed to a design that contains the five mood cues (the mood-optimised design) will have a higher their purchase intention than users who are exposed to a design that does not contain the five mood cues (the original design).

Integration of the Mood Cues into the Design

Given that it was not possible to test all pages on Imerco's website, the mood-optimised design consisted of five pages: The front page, a product overview of vases, a product page for a vase, the shopping cart page and the checkout page. These were selected because they represent a natural shopping flow for a regular user[64]. Moreover, the product overview of vases and the product page for a vase were selected because they have a large number of visitors each month, but low sales rates and a high bounce rate, which

[64] A natural shopping flow for a regular user is: Commencing at the front page, choosing the product category you are interested in, choosing the specific product you are interested in and adding it to the shopping cart, going to the shopping cart and then to the checkout page.

means that a large part of the users who visit these pages leave the website without further interactions[65]. Optimising these pages to make the users interact more and buy the products is therefore of value to Imerco. In the mood-optimised design, the pages in the shopping flow were amended[66] to naturally include the five mood cues. Below, the original product overview page is presented above the mood-optimised product overview page (Figure 2 and 3). Due to lack of space, the remaining four pages in the shopping flow are not presented here, but can be found in the appendix (Appx. 9).

Figure 2: The original product overview page at Imerco's website.

Figure 3: The 'mood-optimised' product overview page.

[65] This insight was drawn from Imerco's Google Analytics account.

[66] The tool Visual Website Optimiser[66] was used for editing the design.

7. Discussion and Implementation of the Qualitative Findings

In this section, the implications of the qualitative findings are discussed. One by one, the findings about each mood cue are discussed followed by an explanation of how that cue was integrated into the mood-optimised design.

7.1. Photos

7.1.1. Discussion

The finding that the motives have to look natural and unstaged can be related to the attribution effect found by Clore and Parott (1994) and Siemer and Reisenzein (1998). As previously mentioned (c.f. section 4.2.3.) they found that when subjects are aware of the source of their mood, affect infusion does not happen, at least not to the same extent. In a similar way, the present findings suggest that in commercial settings such as a webshop, if the subjects are aware that the purpose of the design is to affect their mood, the design does not in fact affect their mood.

The fact that it was difficult to find photos that looked unstaged but at the same time looked neat and displayed a 'dream setting' revealed a paradox: While users on one hand like to be carried away by a dream depicted by a photo, they also become annoyed if the commercial purpose is obvious. How do we identify the 'happy mean' between a stock photo that looks obviously staged and is therefore not trustworthy and a boring- and everyday-looking photo that will not allow the users to dream? This should be further investigated.

The finding from the first focus group that informants find photos of people more appealing is in line with Cyr, Head, Larios, and Pan's findings (2009) that human images with facial features induce the user to perceive a website as more appealing, having warmth and social presence. However, because the two focus groups disagreed on this issue, this issue should be addressed in future research. The contradictory findings may just indicate that there is more to this cue than just including or excluding people from the photos and that there is a fine balance between choosing the 'right' and 'wrong' types of people.

The variations in individual reactions to photos in general shed light on another paradox: While photos clearly cause strong affectionate reactions, it can be difficult to predict the valence of those reactions, and there is a fine line between choosing the right and wrong photo. Some users have negative reactions that are difficult to predict because everyone has individual preferences and different associations with the same photos. Because of individual variance in the informants' reactions, it is difficult to choose a design that affects all users' mood in the same way. In the present research, the mood-optimised design was

based on the mood cues that the majority of the informants agreed upon, but in future research individual differences should be explored further.

Furthermore, there may be cultural differences in the way people react on different designs. It was found that the calm and comfortable look that users like in photos relate to the Danish concept of 'hygge'. In Denmark 'hygge' is something that people seek in their daily lives[67]. Since other countries do not have the same cultural concept, they may react differently to these photos. This should be addressed in future research.

7.1.2. Design Implications

To integrate the findings, a number of photos, including a background photo, were added to the five pages of the shopping flow (Appx. 9). These were all calm because they used subtle colours and only displayed a few neatly placed items. Most of them had a 'hyggeligt' (cosy) look, e.g. by including candle lights or comfortable furniture in the background. They all displayed decorative, trendy-looking and minimalistic home settings. To ensure familiarity, only photos that looked like they had been shot in a regular Danish home were included. None of the photos depicted people. It was difficult to find photos that did not look staged, but at least none of them was obvious stock photos.

Front page (Appx. 9, p. 2): In the top of the original front page, there was a product photo with a strong red background colour and a photo of Imerco's catalogues in red with a grey background. In the mood-optimised design, these were replaced by two similar photos that only used muted grey, light blue and light golden colours. The descriptive texts "offers of the week" and "Christmas catalogues" were added beneath the photos to make them more noticeable. Furthermore, the display of random products on sale on the original front page (a fluency issue mentioned in section 6.1.3.1.) was replaced with large photos symbolising the various product categories. These were accompanied by texts that described the categories.

Product overview page (Appx. 9, p. 3): A large high quality photo in subtle colours was added to the top of the page to create a cosy and tranquil atmosphere. It displays a vase neatly placed on a living room table next to lit candles and comfortable looking couches with pillows in a trendy-looking Danish living room setting.

[67] For example, Danes usually want their homes to look 'hyggeligt' and to create a warm and comfortable ambiance where they can have a good time together with friends and family.

7.2. Background Colour

7.2.1. Discussion

The findings suggest that users react more positively to calm and light background colours that create a contrast between front and background. Yet, individual differences in preference were also found when it comes to colour: Even though the informants generally agreed upon preferring light background colours, two informants were drawn to the strong petroleum blue background. There may be cultural differences on colour preference as well, since the finding that people prefer subtle and light colours may be related to the minimalistic style that is imbued in most Scandinavian design. This should be further investigated.

7.2.2. Design Implications

In the original design, all information was displayed on a plain white background, which made it difficult to process the information because it blended into the background. This was changed by adding a light grey background disc to all pages to create a clearer contrast (Appx. 9). This made all information stand out more clearly.

7.3. Fluency

7.3.1. Discussion

The finding that fluency affects users' mood is interesting because according to theory on metacognitive experiences, which builds on the Feelings-as-Information model, fluency affects judgment in a fashion similar to mood. It has been found that the easier something is to process, the higher the truth level is perceived (Reber & Schwarz, 1999). Further, it has been found that people are less likely to detect an error when the information is easy to process (Song & Schwarz, 2008). According to theory on metacognitive experiences, the reason for this is that people rely on their subjective processing experiences with the information that they base their decision on (Schwarz et al., 2007). Schwarz defines metacognitive experiences as a variety of subjective experiental information or *cognitive feelings* that people use in forming a judgment (ibid.). *Accessibility*[68] and *processing fluency* are those that have been of most interest to researchers (ibid.). Fluency is defined as the ease with which new, external information can be processed (ibid.). The finding that a design that is easy to process affects the users' mood positively fits well with the theory of metacognitive experiences. According to this theory, fluency creates a feeling

[68] Accessibility has to do with how easy it is to recall and generate thoughts about information. It has been found that people judge information as more frequent and probable when it is easy (high accessibility) rather than difficult (low accessibility) to recall examples of it (Tversky & Kahneman, 1973). Further, high accessibility makes people hold an attitude with more confidence (Berger & Mitchell, 1989).

of things going smoothly. When making a decision, people rely on this subjective experience for making their decision, in the same way that they also do with mood (ibid.). This explains why the informants preferred fluent information.

The finding that users prefer large text and image sizes because they are easier to process is in line with Park et al.'s finding that image size in online clothes sales positively affect users' mood. Like present findings, their finding may also be explained by fluency theory.

7.3.2. Design Implications

To integrate the findings in the mood-optimised design, all text and photos were enlarged. However, sales labels on the product overview and product description page were made slightly smaller because the informants found their prominence distracting (however, given their strong red colour and still relatively large size, they were still very noticeable). The light grey background discs further made the information on the website look clearer and more prominent. Product names were shortened because they were long and complicated[69] (Appx. 9, p. 3-4). The product descriptions were rewritten to not include any 'nonsense'[70] (Appx. 9, p. 4) and formatted in two short and precise bullet form presentations. A large amount of unnecessary information was removed from the pages. These are provided in Table 5.

[69] The original product names included size and quantity information (e.g. "Vase Kvist 16.5 cm 1 stk.").

[70] For example, the text "even a simple flower deserves a beautiful vase" (translated from Danish) was deleted.

Unnecessary information that was not included in the mood-optimised design

- **Random products on sale (front page):** Since users find these unnecessary and disturbing, they were removed and replaced with category photos (Appx. 9, p. 2).
- **Category description (product overview page):** The category description was removed, because it fit with the users' description of unnecessary nonsense, e.g. by stating that "lovely flowers only deserve the best vases from your favorite design" (Appx. 9, p. 3).
- **Duplicate 'next page' icons (product overview page):** Since this icon was placed in both the top and bottom of the page, it was removed from the top (Appx. 9, p. 3).
- **Product status (product description page):** The information 'status: in store' was removed because the users find this information unnecessary and annoying (Appx. 9, p. 4).
- **Interaction buttons (product description page):** These include social share buttons and the 'add to wish list' button, which were removed because the users find them disturbing and unnecessary (Appx. 9, p. 4).
- **EAN (bar code) numbers (product description page/shopping cart/checkout page):** This number was removed since the users do not need them and do not know what they are (Appx. 9, p. 4-6).
- **Duplicate 'Go to checkout' buttons (shopping cart page):** Since there were two of these buttons placed close to each other, one of them was removed (Appx. 9, p. 5).
- **Navigation information (shopping cart page):** The information "click on the button 'go to checkout' to check out" was removed, because this information was already written on the actual button (Appx. 9, p. 5).
- **Discount code field (shopping cart page):** The 'add a discount code' field was removed, since it was unnecessary at this stage and only annoyed the users (Appx. 9, p. 5).
- **Duplicate billing information (checkout page):** Since the billing information was displayed twice (in the fields where they are typed in by the user and in another field in the right side of the page), one of them was removed (Appx. 9, p. 6).
- **Duplicate 'back to the shopping cart' icon (checkout page):** Since this icon was placed in both the top and bottom of the page, the latter one was removed (Appx. 9, p. 6).
- **Duplicate shipping information (checkout page)**: Since the box 'delivery and payment address is the same' was checked off, displaying a field for both payment and delivery was unnecessary, and therefore the latter was removed (Appx. 9, p. 6).
- **Login field (checkout page):** Since this possibility is only relevant to users revisiting the website and not to first time users, it was removed (Appx. 9, p. 6).

Table 5: List of unnecessary information that was not included in the mood-optimised design.

7.4. Gift

7.4.1. Discussion

The finding that users generally react positively to receiving a gift when shopping online fits well with Gerrards-Hesse et al.'s (1994) finding that receiving a gift by surprice improves the mood. Yet, individual differences in preference were also found when it comes to receiving a gift, and some informants expressed that it could make them suspicious. Because of this, when using this cue it should be made clear that it is a *free* gift and that it does not oblige the user to pay or commit to anything. In addition, it should be made clear that the gift is a 'special' exception to the regular deal, so users will not expect to receive one each time.

Surprising the customer with a gift at the time of receiving the package is a form of marketing regularly used to form attachments with the users. On the other hand, offering the gift while the customer is still in the webshop is not a widely used practice. Interestingly, the company GoPro actually used this practice in their webshop[71] in December 2015[72], which indicates that it might be an upcoming marketing practice. In this case, the user was informed that a surprise gift would be added to the shipment, but not what the gift was. There may be an advantage in informing the users that they will receive a gift without disclosing the content of it, because the expectation of an unknown gift may improve mood more. This should be investigated in future studies.

7.4.2. Design Implications

To integrate the gift cue, a high quality paring knife (an item that most people can use) was added to the shopping cart with the headline "A Christmas gift from us: Raadvad paring knife" and a short description wishing the user merry Christmas and explaining that Imerco rewards their loyal customers with a free gift in December (Appx. 9, p. 5). "00.00" was stated instead of the price per piece and "Free gift'" was stated instead of the total price to make it clear that the user does not have to pay anything for the gift.

7.5. Advantages

7.5.1. Discussion

Communicating service and product advantages is not a new marketing practice, but the finding that they can influence mood is. This mood cue can be compared to the feedback MIP (cf. section 4.5.2.2), since it

[71] https://gopro.com/

[72] This was discovered by chance, as a friend of mine did his shopping there.

entails giving the user positive feedback on successfully completing a task. According to Clore et al. (2001), people constantly stream their surroundings for affective cues in the form of feedback about progress toward minor sub goals. The service and product advantages communicate to users that some of their online goals have been reached, e.g. finding good offers and a website with free shipping.

7.5.2. Design Implications

A bullet form summary of the advantages of buying vases at Imerco's webshop was created and displayed in the top middle of the product overview page (the 'bullets' are small green checkmarks that signal that they are positive advantages) (Appx. 9, p. 3). The text was made large and clear and placed in the top of the page to make it more noticeable. Due to design limitations, the customer ratings could not be included in this element.

The next step of this research was to test the mood-optimised design quantitatively to measure the effect of the design on purchase intention, product evaluation, website design evaluation and processing style and to test whether this effect is mediated by mood. The methodology used for this will be presented in the following section.

8. Quantitative Methodology

The aim of the quantitative study was to confirm causal links between the mood-optimised design (the independent variable – henceforth the 'IV') and product evaluation, purchase intention, website design evaluation and processing style (the dependent variables – henceforth the 'DVs') and demonstrate that these relationships are mediated by mood. This was tested by means of an experiment because it is the "gold standard" against which the rigor of other strategies is assessed due to its high level of internal validity (Saunders et al., 2009, p. 141). By holding all other variables constant, an experiment makes it possible to provide clear evidence that the treatment caused the observed outcome (Trochim & Donnelly, 2008). In the following, design, procedure and instruments are accounted for.

8.1. Design

To isolate the design variable from other potential causes of the outcome and thereby rule out alternative explanations, a two-group design was employed by assigning the subjects[73] randomly to either a treatment group (presented with the mood-optimised design) or a control group (presented with the original design)

[73] The terms 'subjects' and 'participants' are used interchangeably to refer to the people who took part in the experiment.

(Trochim & Donnelly, 2008). This was necessary to confirm that the DVs have different values when the design is used and when the design is not used.

However, there was a risk that variance between the two groups on gender, age and education level could affect the outcome, since these variables have been found to moderate effects in similar experimental research[74]. Furthermore, there was a risk of variance due to different levels of online shopping experience. To account for variations in these background variables, the study used a randomised block design by dividing the sample into 16 relatively homogeneous subgroups based on the variables (listed in Table 6). Random assignment to experimental groups thereby took place within each block. Because the variability within each block is less than in the entire sample, this reduced the risk of variance due to random uncontrollable factors (noise) and thereby made the predicted relationships between the variables (signal) more clear (Trochim & Donnelly, 2008).

		Males		Females	
		Low education	High education	Low education	High education
Under 45	Shop online once a month or more	Males under 45 with low education who shop online once a month or more.	Males under 45 with high education who shop online once a month or more.	Females under 45 with low education who shop online once a month or more.	Females under 45 with high education who shop online once a month or more.
	Shop online less than once a month	Males under 45 with low education who shop online less than once a month.	Males under 45 with high education who shop online less than once a month.	Females under 45 with low education who shop online less than once a month.	Females under 45 with high education who shop online less than once a month.
Over 45	Shop online once a month or more	Males over 45 with low education who shop online once a month or more.	Males over 45 with high education who shop online once a month or more.	Females over 45 with low education who shop online once a month or more.	Females over 45 with high education who shop online once a month or more.
	Shop online less than once a month	Males over 45 with low education who shop online less than once a month.	Males over 45 with high education who shop online less than once a month.	Females over 45 with low education who shop online less than once a month.	Females over 45 with high education who shop online less than once a month.

Table 6: The 16 subgroups in the randomised block design.

[74] For example, older consumers rely more on heuristic processing (Yoon, 1997), men and women differ in their affectionate reactions and purchase intentions towards advertisements (Putrevu, 2004) and education level moderates technology adoption (Abu-Shanab, 2011).

8.2. Procedure

Subjects were recruited by use of convenience sampling in exchange for a chance of winning two movie tickets. Invitations were distributed to friends and acquaintances on Facebook and by e-mail (Appx. 10.1). Participants signed up by clicking on a link and filling out a short questionnaire (Appx. 10.2) with demographic information that was used for allocating them into blocks. Within each block, participants were randomly assigned a number by use of the tool Random.org (RANDOM.ORG, 2016a). All subjects with equal numbers were assigned to the control group and all unequal numbers were assigned to the treatment group[75].

The data was gathered by means of a survey-experiment. This reduced the participation and administration time, increased the convenience of participating (which increased the sample size) and made the online shopping scenario appear more realistic. Participants were informed that the survey concerned online shopping behavior, but not that it was an experiment.

One week after signing up, the participants received an e-mail (Appx. 11.1) containing a short guide and a trackable link[76] to the survey. Three days later, those who had not filled out the survey received a reminder (Appx. 11.2). The content was identical in all of the e-mails, but the links led to either the treatment or control version of the survey.

The survey (Appx. 11.3 & 11.4) consisted of an online shopping scenario followed by a questionnaire. On the first page, subjects were instructed on how to take the survey. Subsequently, the actual stimulation came in form of an online shopping scenario. Like other researchers (Park et al., 2005), this study used a mock-up website consisting of screen shots of the five website pages included in the shopping scenario (cf. section 6.3.) – either the original (control) or the mood-optimised (treatment) version. They were guided through each page and asked to imagine being in a real shopping situation and use as much time as they would normally do. Inspired by Swinyard (1993), the screen shots were accompanied by written scenarios[77] to simulate a personally relevant situation (Appx. 11.3, p. 5 & 11.4, p. 5) because various researchers have found that such written descriptions increase behavioral consistency (Graham et al. 1989, Martin 1991, Merluzzi and Biever 1987 and Fazio and Zanna 1981 as cited in Swinyard, 1993). To minimise the risk of random factors, all factors beyond the design were held constant. After exposure to

[75] In blocks with an unequal number of subjects, the subject with the highest number was allocated to a new group. Subjects within this last group were then randomly assigned to the two experimental groups.

[76] A link that made it possible to track whether the person had completed the survey.

[77] E.g. "Imagine that you decide to add five vases of this type to the shopping cart. When you reach the shopping cart it looks like this".

the shopping scenario, participants completed a questionnaire that measured the DVs and collected demographic information.

8.3. Instruments

All instruments used for measuring the variables had been tested by other researchers. Therefore, pretests were not necessary. The instruments are accounted for in the following.

8.3.1. Mood Measurements

Most researchers use at least one self-report measurements to report on mood (Westermann et al., 1996). The present research only used one self-report measurement, since a large number of mood questions would increase the risk of demand effects[78] (Park et al., 2005; Westermann et al., 1996). Participants were asked to rate their current mood on a 10-point numeric rating scale with the end categories 'Really good' (1) and 'Really bad' (10) (Saunders et al., 2009). By comparing effect sizes of various mood studies, R. J. Larsen and L. M. Sinnett (1991) has found that effect sizes are generally larger when a self-report measure is used for measuring mood. They argue that the self-report measure is the most valid measurement. Accordingly, the present research regards this as the strongest mood-measuring instrument. Still, asking the subjects to rate their own mood causes some validity issues, since it can be difficult to evaluate one's own mood (Olsen, 1998), and thus there exists a risk of some respondents just choosing a random answer. In addition, it is a common provision in mood research to measure mood with more than one instrument to increase the chance of detecting an effect (Larsen & Sinnett, 1991). Accordingly, two additional non-self-report instruments were used. Subjects were blind to the purpose of both.

The first was the *word-pleasantness* method, which is based on the theory that subjects in a good mood rate unfamiliar words higher than subjects in a bad mood do. It has been used previously by A.M. Isen, M. M. Johnson, E. Mertz, and G. F. Robinson (1985) and was replicated by C. Verheyen and A.S. Göritz (2009) who found that their subjects' ratings of the unfamiliar word saté reflected their self-reported mood (Isen, Johnson, Mertz, & Robinson, 1985; Verheyen & Goritz, 2009). In addition, Forgas (1995) argues based on a literature review that the impact of mood on judgment is stronger the more unfamiliar the stimulus is. In the present study, subjects were therefore asked to rate eight unfamiliar words on a ten-point numeric bipolar scale with the end categories 'Very unpleasant' (1) and 'Very pleasant' (10). The words were selected from a study of unfamiliar Danish words according to which the words *ætanol, assimilation, kulance, endossere, kondemneret, stipuleret, akklamation, approbere, kumulere* and

[78] Subjects guessing the purpose of the study.

allokering are unfamiliar words that most Danes do know the meaning of (Dansk & Sprognævn, 1985). Since the study is from 1985, it is possible that this has changed. When pilot testing the questionnaire, the informants were therefore asked to state whether they knew any of the words on the list. Since some of them knew *ætanol, allokering* and/or *assimilation*, these were dropped from the list.

Since several researchers have found that people in a positive mood make decisions more quickly than people in a negative mood (Duque, Turla, & Evangelista, 2013; Velten, 1968), a second non-self-report measure was employed to measure the time it took subjects to process and answer the questionnaire. Since no other researchers had measured time in an online survey, a new method was invented for this. The survey tool used (SurveyXact) did not have a time measurement function, so the participants were instead asked to note down the current time in the format 00:00 (e.g. 23:41) just before and just after finishing the DVs in order to be able to measure the completion time. To avoid demand effects, they were not made aware of the real purpose of this. The total time was calculated by subtracting the two numbers.

8.3.2. Processing Style Measurements

As noted earlier, heuristic processing involves more effortless and uncritical thinking than systematic processing (Chaiken & Ledgerwood, 2007; Kahneman, 2002). From this follows that people who use heuristic processing remember less of what they have processed. For example, F. I. Craik and T. Endel (1995) found that people who use heuristic processing can remember fewer words from a list[79] and J. Saegert & R. Young (1981) found that people who use heuristic processing can recall less brand names from magazine ads[80]. Forgas (1995) also argues in his Affect Infusion Model that processing styles are distinguishable through the analysis of memory. Based on these findings, memory is considered a valid measurement for processing style. The present study therefore measured processing style by average performance on four questions[81] that tested how much subjects could remember about what they had seen in the online shopping scenario. In order to avoid annoying the participants, the questions were not presented as a test but just general measurements.

[79] Craik & Endel manipulated processing style by asking questions (before exposure to the words) that were designed to stimulate either heuristic processing (questions about the word's physical appearance) or systematic processing (questions about the word's meaning) (Craik & Endel, 1975).

[80] Saegert & Young manipulated processing style by playing tape-recorded questions at the time of exposure to the ads. The questions were either designed to stimulate heuristic processing (e.g., "Is the brand name in blue letters?") or systematic processing (e.g., "Have you heard of this brand name before?") (Saegert & Young, 1981).

[81] The questions all concerned a vase they had seen in the scenario: the name, the delivery time and the price before and after discount.

8.3.3. Measurements for the Remaining Dependent Variables

The remaining DVs were measured by means of 10-point numeric bipolar rating scales. Purchase intention was measured on a likelihood scale since this is the most commonly used instrument for researchers examining the relationship between mood and purchase intention (Donovan & Rossiter, 1982; Park et al., 2005; Swinyard, 1993). To increase the chance of detecting differences between the two experimental groups, two variations of this measure were used: One measuring their purchase intention for a specific vase they had been presented with in the scenario (henceforth 'specific purchase intention') and one measuring their general purchase intention for any product on the website (henceforth 'general purchase intention'). In both cases the participants were asked to rate the likelihood that they would buy the/a product on a scale with the end categories 'Very unlikely' (1) and 'Very likely' (10).

Inspired by the procedure used in Gorn et al.'s study of mood and product evaluation (Gorn et al., 1993), product evaluation and website design evaluation were measured by means of 10-point scales with the end categories 'Like it a lot' (1) and 'Do not like it at all'.

On the last page, the following demographic variables were provided: Gender, age, postcode, highest school degree obtained and how often they shop online. The last variable was used to sort out the ones who had never shopped online.

8.4. Data Types

In order to make correct calculations with the data (in the program SPSS), data types were defined for each variable. Those that were impossible to define numerically or rank because they represent different categories that do not imply any order were defined as *nominal* (Greasley, 2008; Saunders et al., 2009). All variables that took the form of quantifiable numbers that go from high to low with the same mathematical span between them were given the data type *scale* (ibid.). The coding book can be found in the appendix (Appx. 12.10).

8.5. The Questionnaire

In questionnaires, validity is defined as the ability of the questionnaire to measure what the researcher intends it to measure (Saunders et al., 2009). The questionnaire was designed to enhance validity by ensuring that all respondents understood the questions in the way as intended by the researcher (Saunders et al., 2009). The risk of misunderstandings was reduced by using short sentences and avoiding linguistic and grammatical errors and by using words that are used in everyday speech (Olsen, 1998). In addition,

the questionnaire was pilot tested on five subjects in iterative cycles. According to Jacob Nielsen (2011), five subjects allow you to discover most usability problems (Nielsen, 2011).

The participants were blind to the purpose of the study in order to reduce bias. The questions were placed in an order that reduced the risk of questions affecting subjects' performance in subsequent questions (Cramer & Howitt, 2004). Optimally, mood should be measured as the first variable after the shopping scenario to avoid that the subjects' mood changes as the other DVs are completed. However, this conflicts with the goal of avoiding demand effects: If participants are asked to rate their mood directly before the DVs, they can easily guess the purpose of the study, and this can affect their subsequent performance. Consequently, product evaluation, specific/general purchase intention and website design evaluation were measured directly after the shopping scenario, followed by the self-report and the word-pleasantness ratings mood measurements.

It could have been beneficial to measure mood both before and after exposure to the design (comparing effect sizes instead), since this would reduce the risk of validity issues due to uncontrollable factors that may affect the subjects' baseline mood (e.g. the weather). Furthermore, it would have made it possible to conclude whether mood differences between the two experimental groups are due to the original design affecting the users' mood negatively, the mood-optimised design affecting the mood positively, or both. Since mood is only measured once, it can only be concluded whether or not the mood-optimised design affects mood more positively than the original design. Still, as previously mentioned (cf. section 4.2.2.) the mood effect does not happen if the subject is aware of the source of their mood (Clore et al., 2001; Forgas, 1995; Schwarz, 2011). Since placing the mood question both before and after the shopping scenario could make the subjects aware of the purpose (measuring the effect of the design on mood) and thereby aware of the source of their mood, there was a danger that affect infusion would not take place. That is why mood was only measured after exposure.

The memory questions (measuring the last DV: processing style) were placed after the mood measurements because researchers have found that performance on tests affect mood (the feedback MIP) (Isen & Shalker, 1982; Swinyard, 1993). Since the subjects were blind to this measure, the risk of demand effects was small. Lastly, the demographic variables were measured.

8.6. Hypothesis Testing

Because the hypotheses predict that product evaluation, general and specific purchase intention, website design evaluation and processing style are dependent on the website design, they are all classified as DVs

(Y) while the design is the IV (X) (Greasley, 2008). Since mood is the reason why X is expected to influence Y, mood is classified as the mediator (M) which is a variable that accounts for the relation between the predictor and the criterion by intervening between stimulus and response, thereby specifying how or why the effect occurs (Baron & Kenny, 1986). The conceptual model is provided in Figure 4.

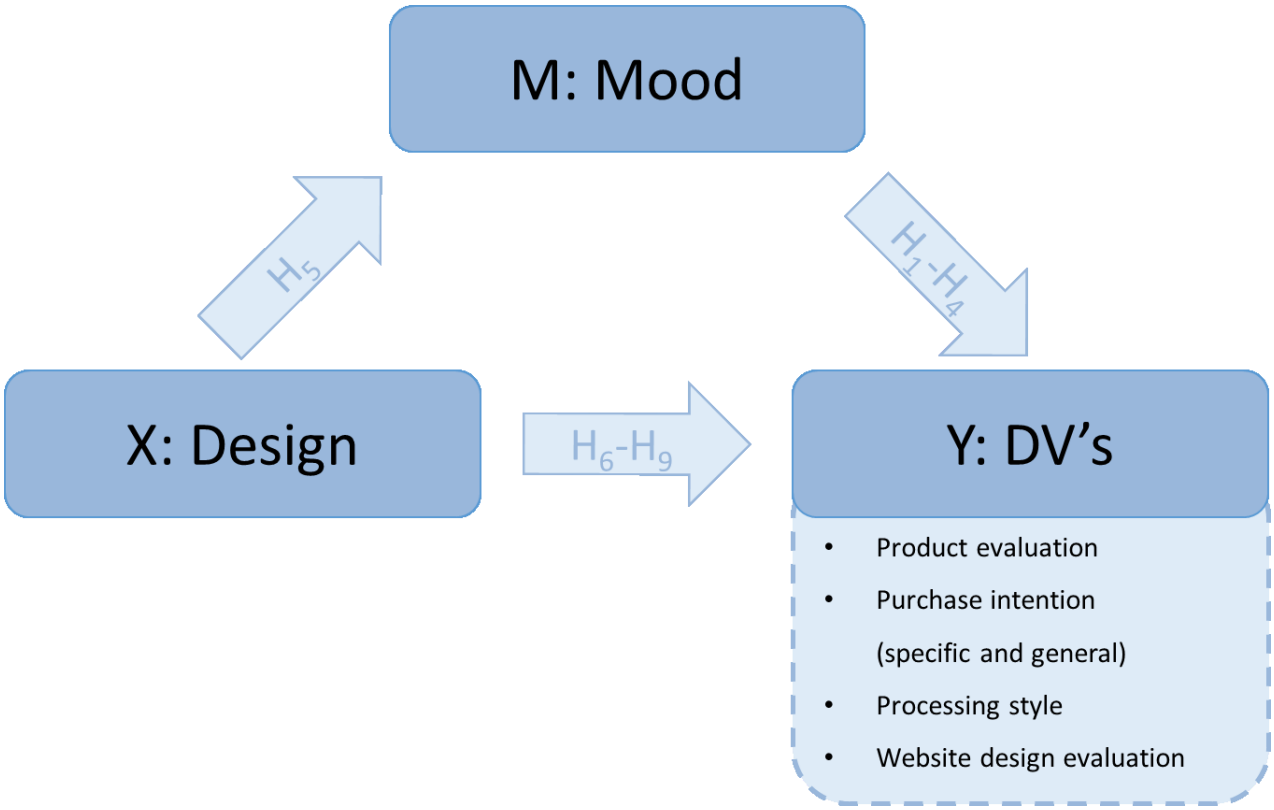

Figure 4: The conceptual model for the effect of the design on the dependent variables.

For each expected relationship between two variables, a null hypothesis was formed to state the absence of a relationship between them, which could be either retained or rejected in the data analyses (Greasley, 2008). In a null hypothesis, the hypothesis is reversed to the contrary in order to test the probability that there is a causality between the two variables (Andersen & Hansen, 2000). For each null hypothesis (H0) an alternative hypothesis (H1) was formed. These are stated in Table 7.

Null and alternative hypotheses

Mood and website design evaluation

$H0_1$: There is *no relationship* between mood and website design evaluation.

$H1_1$: There is a *positive relationship* between mood and website design evaluation.

Mood and product evaluation

$H0_2$: There is *no relationship* between mood and product evaluation.

$H1_2$: There is a *positive relationship* between mood and product evaluation.

Mood and processing style

$H0_3$: There is *no relationship* between mood and processing style.

$H1_3$: The better the mood, the less systematic the processing style, which means that there is a *negative relationship* between mood and memory score.

Mood and purchase intention

$H0_4A$: There is *no relationship* between mood and specific purchase intention.

$H1_4A$: There is a *positive relationship* between mood and specific purchase intention.

$H0_4B$: There is *no relationship* between mood and general purchase intention.

$H1_4B$: There is a *positive relationship* between mood and general purchase intention.

Design and mood

$H0_5$: Users who are exposed to the mood-optimised design will not be in a different mood than users exposed to the original design.

$H1_5$: Users who are exposed to the mood-optimised design will be in a *better mood* than users exposed to the original design.

Design and website design evaluation:

$H0_6$: Users who are exposed to the mood-optimised design will not evaluate the design of that website differently from users exposed to the original design.

$H1_6$: Users who are exposed to the mood-optimised design will evaluate the design of that website *more positively* than users exposed to the original design.

Design and product evaluation:

$H0_7$: Users who are exposed to the mood-optimised design will not evaluate a product from the website differently from users exposed to the original design**.**

$H1_7$: Users who are exposed to the mood-optimised design will evaluate a product from the website *more positively* than users exposed to the original design.

Design and processing style:	
$H0_8$:	Users who are exposed to the mood-optimised design will not use a different processing style from users exposed to the original design, which means that there will not be a difference in how much the two groups can remember.
$H1_8$:	Users who are exposed to the mood-optimised design will use a *less systematic (more heuristic) processing style* than users exposed to the original design, which means that they will *remember less of what they saw on the website.*
Design and purchase intention:	
$H0_9$A:	Users who are exposed to the mood-optimised design will not have a different specific purchase intention for buying the vase from users exposed to the original design**.**
$H1_9$A:	Users who are exposed to the mood-optimised design will have a *higher specific purchase intention* for buying the vase than users exposed to the original design.
$H0_9$B:	Users who are exposed to the mood-optimised design will not have a different general purchase intention for buying the vase from users exposed to the original design**.**
$H1_9$B:	Users who are exposed to the mood-optimised design will have a higher *general purchase intention* for buying the vase than users exposed to the original design.

Table 7: Null and alternative hypotheses.

8.7. The Sample

214 participants originally signed up for the survey, but only 155 completed the survey[82]. Since there was only one participant in the age group 65+, this participant was discarded. Additionally, four participants were discarded because their calculated finishing time was unrealistically low or high. After filtering the data, there were 73 respondents left in the treatment group and 75 in the control group. In order to have the same number of participants in each group, two numbers between 1 and 75 were randomly generated using Random.org (RANDOM.ORG, 2016b). The participants with the corresponding numbers were discarded. Of the 146 remaining participants, 87 (59.6%) were female and 59 (40.4%) were male. The average age was 35.34 (SD=11.363) (Appx. 12.2).

[82] 77 people completed the survey in the control group and 78 people completed the survey in the treatment group.

8.7.1. Sample Size

The optimal sample size for the experiment was calculated by use of Al-Therapy Statistics' sample size calculator (AI-Therapy, 2016). The calculation was based on a confidence (significance) level of 95% and a power level of 80%, which are the statistical standards. The effect size was based on Westermann, Stahl and Hesse's (1996) findings of a mean effect size (r) of 0.41 across 138 mood induction studies. Since the tool calculates sample sizes for use in T-tests and samples have to be 15% larger when testing with Mann Whitney U tests (GraphPad, 2015; IBM, 2010; Simon, 2000), the calculated sample size was increased by 15% to obtain the correct size. Based on this calculation, the sample size should be 219 (Appx. 12.2).

However, the same effect size as in the above-mentioned 138 mood induction studies could not be expected in the present study, because it uses a different mood induction procedure. Whereas the 138 studies were conducted in controlled lab experiments with procedures designed solely for manipulating mood, the present study naturally incorporates the mood manipulation into the design of a website. Thus, the effect size is expected to be much smaller in the present study. In addition, Göritz' (2007) study of mood induced in an online survey demonstrated that people's baseline mood is generally higher in an online context than offline, which suggests that a further mood lift will be smaller. This means that the present study should have used a larger sample to be sure to detect small effect sizes. For example, to confirm an effect of 0.2, a sample size of about 906 (453 in each group) is needed (Appx. 12.2). Since the actual sample size was 146, the power level was only around 0.20 (Appx. 12.2) (more on this issue in the section 10).

8.7.2. External Validity

To make the sample representable it should have been drawn by use of random selection from the population, but unfortunately, Imerco did not want their customers to be contacted with invitations for surveys. Another solution could have been using stratified random sampling to draw a sample that was representative in terms of gender, age, level of education and level of online shopping experience[83]. This would have made the sample generalisable to the population. However, this was not possible due to difficulties of finding participants for the study.

Additionally, there may be a difference between (1) those who saw the invitation, chose to participate and completed the survey, (2) those who saw the invitation, chose to participate and did not complete the

[83] The subgroups for the randomised block design were based on the same variables.

survey and (3) those who saw the invitation and did not choose to participate, which reduces the generalisability.

According to the rough sample size guide provided by Saunders et al., the minimum sample size needed to be able to generalise to the population size (the number of visitors on Imerco's website in 2015[84]), is 384 participants (Saunders et al., 2009). The fact that only 146 subjects were recruited further limits the generalisability.

After gathering the data, it was tested whether the sample was representative of the population in regards to gender and age distribution by use of a one sample T-test with the hypotheses listed in Table 8.

Null and alternative hypotheses for testing the representability	
Gender distribution:	
HO:	The gender distribution in the sample is not different from the gender distribution in the population.
H1:	The gender distribution in the sample is different from the gender distribution in the population.
Age distribution:	
HO:	The age distribution in the sample is not different from the gender distribution in the population.
H1:	The age distribution in the sample is different from the gender distribution in the population.

Table 8: Null and alternative hypotheses for testing the representability of the sample in regards to gender and age.

Both the alternative hypotheses were confirmed. This suggest that in regards to gender the sample value of 1.60 (59.6% women) is not representative (p=0.000) of the population value of 1.782 (78.2% women) and that in regards to age the sample value of 1.19 (19.3% above 45) is not representative (p=0.000) of the population value of 1.46 (46.65% above 45) (Appx. 12.1 & 12.3).

Even though the sample is not representative, it can indicate some tendencies, which have both practical and theoretical implications. These will be discussed in the results section.

[84] The exact number of visitors is confidential and has been drawn from Imerco's Google Analytics account (based on tracking in the period 01-01-2015 to 31-12-2015).

8.8. Limitations of the Experiment

In this section, the credibility of the findings will be discussed in regards to both reliability and validity. While validity (or internal validity) refers to whether the findings are really about what they appear to be about, reliability refers to the extent to which other researchers would obtain the same results using the same measures (Saunders et al., 2009).

There are two possible types of error involved in statistical conclusions: Type 1 and Type 2. Type 2 error refers to a situation where the researcher concludes that there is *no* relationship between variables when in fact there *is* one, and Type 1 error refers to a situation in which the researcher concludes there *is* a relationship when in fact there is *not* (Trochim & Donnelly, 2008). While the odds of saying there is *not* a relationship when in fact there is *none* is defined by the significance level (e.g. 95%), the odds of saying there *is* a relationship when in fact there *is* one (that is, the odds of detecting an effect if one really exists) is defined by statistical power. In the following, the risks of making a Type 1 or a Type 2 error will be discussed in relation to the present research.

8.8.1. Type 1 Error

In the present research, the risk of Type 1 error was minimised by only accepting results that were statistically significant, which means that the achieved results can be expected in 19 out of 20 times (due to the 95% significance level). Furthermore, what has been labeled as "fishing" for results - the practice of performing more and more tests until reaching the conclusion desired by pure chance - was avoided (Trochim & Donnelly, 2008, p. 255).

8.8.2. Type 2 Error

According to Trochim and Donnely (2008), most conclusion validity issues have to do with Type 2 errors, since it can be difficult to "find the needle in the haystack" (Trochim & Donnelly, 2008, p. 254), that is, to identify relationships in data in the first place. Type 2 error assumes that a relationship has already been demonstrated and is therefore only concerned with whether that relationship is causal (the difference could have been caused by another factor). Type 2 error has to do with effect size, that is, how strong the effect is in relation to the noise of random factors that also influence the results and thereby make the effect harder to see (ibid.). Since Type 2 errors are caused by low reliability of measures or low validity due to variance in the two groups, these will be discussed in the following.

8.8.2.1. Reliability of Measures

The method and procedure used to collect and analyse the data has been thoroughly described to make it possible for other researchers to replicate the study. As discussed in section 8.3. and 8.5., the risk of subject bias[85] was reduced by making the subjects blind to the purpose of the study and sorting the questions in a way that would minimise the risk of demand effects. However, the use of a survey design introduced a threat to reliability due to subject error[86] because the subjects participated at various times. This is related to the risk of variance between the groups due to uncontrollable factors (such as the time at which the subjects participated), which will be discussed next.

8.8.2.2. Validity Issues Due to Experimental Design

In experiments, validity is defined as the extent to which the findings can be attributed to the interventions rather than any flaws in the research design (Saunders et al., 2009). Using a randomised block design made it possible to reduce noise of random factors exerting an influence on the results. However, using a survey design for collecting the data made it difficult to control variance caused by environmental variation such as the current time, the weather, individual life events, or distractions in the background of the computer screen, which may have affected the subjects' answers. Another threat to validity is the relative heterogeneity of the sample (due to the use of convenience sampling), since group members are likely to vary more on measures if the sample is diverse (Trochim & Donnelly, 2008). The risk of variance between the two groups could be lowered by conducting the experiment in a lab instead of a survey, which increases the level of control, but this would make it difficult to obtain a large sample.

These issues raise the risk of committing Type 2 error, which raises the odds of saying that there is no effect when in fact there is one. It is thereby possible that the present study was not able to detect the effects because of all the noise from uncontrollable variables or unreliable measures.

8.8.3. Power

Power is the odds of saying there is a relationship when in fact there is one – that is, the odds of detecting an effect if one really exists. Power is important because it reflects the degree to which we can detect the differences we expect: The smaller the expected effect size, the larger the sample has to be. To reach valid conclusions, the goal is to achieve a balance between sample size, effect size, significance level, and

[85] Subject bias: Participants giving the response they think the researcher wants or what they think is the correct answer (Gratton & Jones, 2014).

[86] Subject error: Subjects responding differently depending on when they are asked to supply the data (Saunders et al., 2009).

power. This balance should allow the maximum level of power (the statistical standard is 0.80) to detect an effect in the sample. As previously argued, the effect size was expected to be low (cf. section 8.7.1.). Thus, a larger sample was needed to be able to confirm the results with a high level of statistical power while retaining the significance level at 95%. As mentioned earlier, for a sample size of 146 subjects and an effect size of 0.2, the power level is only 0.20 (Appx. 12.2). This means the study would only be able to detect a statistically significant effect in one out of five times.

Due to the high risk of Type 2 error and low level of statistical power, there is a high risk that this study has concluded that there is no difference in instances where in fact there is one. This would have been avoided if the sample size had been larger and the reliability and validity had been higher.

8.8.4. Validity Issues Due to Instrumentation

To increase construct validity[87], the present research uses measures that have been successfully used by other researchers. Still, they had to be adapted to a new context, since no researchers had used these measures in an online survey experiment before, and this may have introduced some validity issues. For example, the way that time was measured was not optimal because it only measured minutes and not seconds, which made the results imprecise and limited the extent to which the two experimental groups could be compared. Future studies should address this issue by finding a more suitable way of measuring time.

After analysing the results, it has become clear that it would have made better sense to measure mood and DVs *during* the online shopping scenario instead of after, since the purpose of the design is precisely to affect the users' mood and decisions *while* they are shopping. It was not the purpose of the design to affect judgments made by the users after leaving the site, even though product evaluations and purchase decisions made during the visit might also affect the users' behavior after the visit (e.g. by buying the products in the offline store or by revisiting the shop later to make a purchase). The only variable that was measured *during* the shopping scenario was processing style. Even though the actual memory questions designed to measure processing style were not asked until after the shopping scenario, the memory score measures the level of systematic processing used during the shopping scenario, and this makes the processing style measure the most reliable one.

[87] Construct validity: The extent to which the measurements actually measure the presence of the constructs they are supposed to measure (Saunders et al., 2009).

8.9. Choice of Statistical Tests

8.9.1. Comparing Groups for Statistical Differences

To test the causal effect of the IV on the DVs and the mediator, it was necessary to compare the two experimental groups for statistical differences. To do this, either a t-test or a Mann-Whitney U test was needed (Greasley, 2008).

A t-test, which is the most sensitive in discerning significant differences in two scores, can only be used on parametric data. This means that the level of measurement has to be interval or ratio, the scores approximate a normal distribution[88] and the variance of scores in both groups be relatively similar (the scores have to follow approximately the same pattern when arranged in a histogram) (Greasley, 2008). The DVs all meet the first demand, but in order to test whether they also meet the second and third demand histograms were constructed. The histograms below illustrate the distribution of specific purchase intention scores in the two groups (Figure 5). The data does not look normally distributed as most of the data is skewed towards the higher end of the scale (ibid.). The patterns of the two histograms are not completely similar either. The same was found for all of the other DVs and the mediator (Appx. 12.4.1).

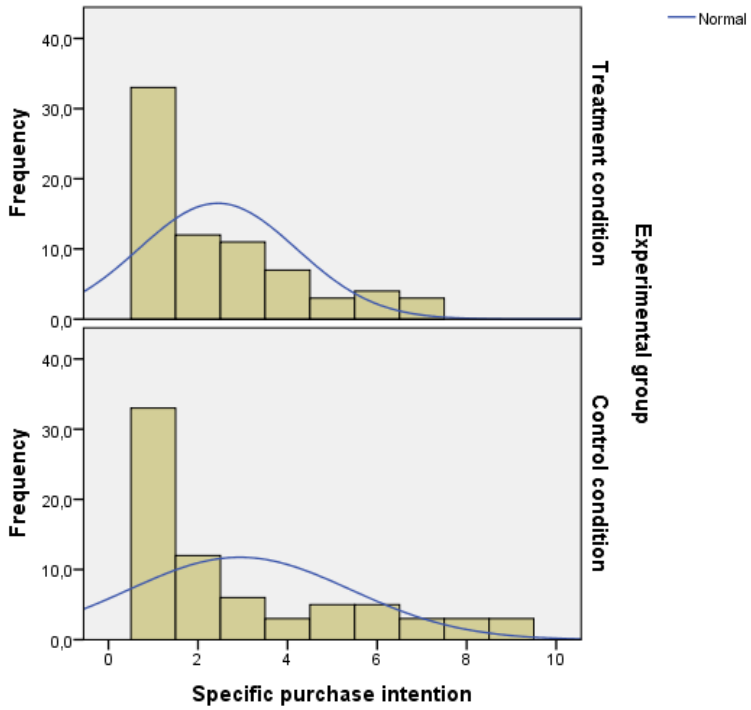

Figure 5: Histograms illustrating the distribution of the specific purchase intention scores in the two groups.

[88] Normal distribution means that the data is distributed equally on both sides of the mean and that the mean and median are identical or close to being identical.

To confirm this indication, a Shapiro-Wilk and a Kolmogorov-Smirnov test was conducted on each of the DVs to test for normality. Since all tests included significant results (p=<0.05) this confirmed the suggestion that the data was not normally distributed (Appx. 12.4.2).

This means that the non-parametric test Mann-Whitney U was the correct one to use[89] (Greasley, 2008). This test involves ranking all scores from both groups in order of magnitude (assigning a rank to each score) and then calculating the mean rank for each of the two groups and the chance that the result occurred by chance (p-value) (ibid.). Because the Mann-Whitney U test was used, medians rather than means were compared in the analysis, since these are more precise when the data is not normally distributed. A median is the midpoint when all the scores are ranked in order of magnitude (ibid.). Means, which are the sum of scores divided by the number of scores, and standard deviation, which is the average deviation from the mean for each group (ibid.), will still be reported on to give the reader the opportunity to see which direction the result would lean towards if the data had been normally distributed. The significance of the results (the likelihood of the results occurring by chance) are measured by the p-value. For the results to be statistically significant, the p-value has to be below 0.05, which means that the chance of the results occurring by chance is below 5% (ibid.).

8.9.2. Correlation Analysis

Since the mood variable was not isolated from all other potential causes of the outcome (like the design variable), it could not be tested whether mood caused the changes in the DVs. Instead, a Pearson correlation analysis (tests for linear correlations between scale variables) was conducted to test whether an increase in mood is associated with an increase in product evaluation, website design evaluation and specific/general purchase intention and with a decrease in memory score (suggesting that a more heuristic processing style was used) (Appx. 12.8). In a Pearson correlation analysis, the strength of a correlation is measured by means of the value r (effect size). If the r-value is negative (e.g. -0.5), the correlation is negative. If the r-value is positive (e.g. +0.5), the correlation is positive. According to the general guideline, the correlation is weak when the r-value is between 0.1 and 0.4 and strong if it is above 0.5,

[89] It should be mentioned that the data distribution of the word ratings of *endossere* was normally distributed in the treatment condition, since neither of the tests showed significant results (Kolmogorov-Smirnov: p=0.006, Shapiro-Wilk: p=0.062). However, since the control condition showed significant results for the same variable (Kolmogorov-Smirnov: p=0.004, Shapiro-Wilk: p=0.004), which means that the data distribution in the two conditions do not follow the same pattern, it should still be tested with the Mann-Whitney U test (Greasley, 2008).

while a value close to zero indicates the absence of a relationship (Greasley, 2008). As with the Mann-Whitney U test, the significance of the results was measured by the p-value.

8.9.3. Mediation Analysis

To test whether there was a mediation R. M. Baron and D. A. Kenny's (1986) approach was used. According to them, for mediation to occur, the following statements have to be true:

X significantly accounts for variations in M

M significantly accounts for variations in Y

X significantly accounts for variations in Y

However, since the tests did not identify any variable for which all three statements were true, the last step in confirming mediation (multiple regression to confirm that X and M predict Y) was dropped (Baron & Kenny, 1986).

8.10. Preliminary Analysis

As previously mentioned, the self-report mood measure instrument is regarded as the strongest mood-measuring instrument. To validate the word-pleasantness instruments, they were therefore held up against the self-report measure in a correlation analysis (Appx. 12.5.1). The following unfamiliar words were positively correlated with the self-report: *Kulance* (r=0.210, p=0.011), *endossere* (r=0.285 p=0.000), *approbere* (r=0.216 p=0.009) and *kumulere* (r=0.180, p=0.030). All though none of these are strong correlations (r=<0.5), this means that an increase in the self-reported mood is associated with an increase in the word-pleasantness rating of these four unfamiliar words. Since the ratings of the words *kondemneret, stipuleret, akklamatio*n, and *allokering* were not positively correlated with the self-report, these were not considered valid instruments for measuring mood. Consequently, these were dropped from the analysis.

As previously stated (cf. section 8.3.2.), processing style was measured by average performance on four memory questions. To calculate the total memory score, each of the four memory variables were recoded into new variables with the values 1 for 'correct answer' and 0 for 'incorrect answer or cannot remember'. The total memory score variable was calculated by adding the four values from the memory-variables together and dividing them by four (Appx. 12.5.2).

9. Quantitative Results

In this section, the results of the statistical tests will be presented. The tests can be found in the appendix (Appx. 12).

9.1. Descriptive Statistics

First, the descriptive results across both experimental groups (Appx. 12.6) will be presented.

The descriptive results suggested a large difference between on one hand the participants' product evaluation and specific purchase intention and on the other hand their general purchase intention and website evaluation. While the median product evaluation (for the vase) was 3 (mean=3.91, SD=2.206) and the median specific purchase intention (for the vase) was 2 (mean=2.68, SD=2.158), the median general purchase intention was 8 (mean=7.04, SD=2.280) and the median website design evaluation was 7 (mean=6.58, SD=1.772). This finding suggests that the participants generally liked the website and its products, but not the vase.

The self-reported mood across both groups was generally high with a median of 8 (mean=7.47, SD=1.632) and most word-pleasantness ratings (non-self-report mood measures) were placed around the medium score[90]. This is in line with the findings of other mood researchers who have found that baseline mood is generally biased in a positive direction (Gerrards-Hesse et al., 1994) as well as with Göritz' (2007) finding that people's baseline mood is generally higher in an online context than offline. It took the participants around three minutes to finish the questions (median=3, mean=3.44, SD=1.270).

The average score for the memory questions was 2.42 (median=3, mean=2.42, SD=1.002), which means that the subjects generally answered two to three out of the four questions correctly. While most subjects (93.8%) correctly remembered the name of the vase, only around half remembered the price before (58.9%) and after (63.7%) discount. Only 26% remembered the delivery time.

[90] Scores for word-pleasantness ratings*: Kulance:* median=5, mean=5.35, SD=2.460. *Endossere:* median=5, mean=4.95, SD=1.963. *Approbere:* median=5, mean=4.41, SD=1.852. *Kumulere:* median=5, mean=5.50, SD=1.959.

9.2. Causal Relationships

In this section, the results from comparing the groups for statistical differences (Appx. 12.7) are presented. Firstly, the results on the effect of the design on the DVs are presented, followed by the results about the effect of the design on mood.

9.2.1. The Design's Effect on the Dependent Variables

The design's effect on the DVs was tested to confirm the overall expectation that the mood-optimised design makes users like a website and its products more, makes them want to buy more and makes them process the website more heuristically.

The difference between the two samples in product evaluation, specific/general purchase intention and website design evaluation was not large enough to suggest significant differences. However, a causal relationship between design and memory was identified, since there was a significant difference between the two samples on memory score. This confirms the prediction that the mood-optimised design makes people use a less systematic/more heuristic processing style. In the following sections, the test results (Appx. 7.1) that led to these conclusions are presented.

9.2.1.1. Product Evaluation

The median product evaluation was 3 in both experimental groups, which means that there was no difference between them. There was a small difference in the two means, which was 3.78 (SD=2.043) for the treatment group and 4.04 (SD=2.365) for the control group, which pointed in the opposite direction of what was hypothesised. Nevertheless, since the data was not normally distributed, the mean was not reliable because the dataset included some extreme values that could distort the mean (which was also reflected in the high standard deviations) (Greasley, 2008). Furthermore, the Mann-Whitney U test did not find this difference to be statistically significant (p=0.652), which means that there was a strong possibility that these scores occurred by chance. Therefore, $H0_7$ could not be rejected.

9.2.1.2. Specific Purchase Intention

The median specific purchase intention (for the vase) was 2 in both groups, which means that there was no difference between them. There was a small difference in the two means which was 2.44 (SD=1.764) for the treatment group and 2.93 (SD=2.480) for the control group. Again the means pointed in the opposite direction of what was hypothesised, but the Mann-Whitney U test did not find this difference to be statistically significant (p=0.551). Therefore, $H0_4$A could not be rejected.

9.2.1.3. General Purchase Intention

The median general purchase intention (for any product on the website) was 8 in both groups, which means that there was no difference between the groups. There was a small difference in the two means which was 7.16 (SD=2.217) for the treatment group and 6.92 (SD=2.350) for the control group. This pointed in the direction of what was hypothesised. However, the Mann-Whitney U test did not find this difference to be statistically significant (p=0.628). Therefore, $H0_9B$ could not be rejected.

9.2.1.4. Website Design Evaluation

The median website design evaluation was 7 in both groups, which means that there was no difference between the groups. The mean was almost the same in the two groups as well, 6.55 (SD=1.826) for the treatment group and 6.62 (SD=1.729) for the control group. The Mann-Whitney U test did not find a significant difference. Therefore, $H0_6$ could not be rejected.

9.2.1.5. Processing Style

As previously mentioned, the level of systematic/heuristic processing involved in processing the online shopping scenario was measured by means of a memory score. The control group had significantly (p=0.036) higher total memory scores[91] (median=3, mean=2.60, SD=0.924) compared to the treatment group (median=3, mean=2.25, SD=1.051)[92], which means that they used a higher level of systematic processing. This meant that $H0_8$ could be rejected and $H1_8$[93] confirmed. To determine the strength of this relationship, the effect size was calculated[94] based on the two means and standard deviations. The effect size was 0.174, which means that there is an effect, but it is weak (r=>0.5) (Becker, 2015; Greasley, 2008). As discussed previously, the chance of detecting such a small effect with a sample size of 146 is less than 20 in 100. This may explain why no effect was found between many of the other variables: If

[91] When the memory scores for each memory question were compared, it was found that the control group performed better on all of the four memory questions, but the difference was largest on their performance on the first question (the name of the vase) and third question (the price of the vase after discount). On the first question, the mean score was 0.90 for the treatment group and 0.97 for the control group (p=0.086) and on the third question the mean score was 0.507 for the treatment group and 0.671 for the control group (p=0.044).

[92] Even though medians were the same, the means were different, and the Mann-Whitney U test, which controls for data that is not normally distributed, found this difference to be statistically significant (the data was slightly skewed towards the lower end and close to being normally distributed, but the Shapiro-Wilk and Kolmogorov-Smirnov tests found that it was skewed) (Appx. 12.4.2). Furthermore, an independent samples T-test also found this difference to be significant (p=0.31) (Appx. 12.7.1). Therefore, the test results are considered valid.

[93] $H1_8$ (as stated in section 8.6.): Users who are exposed to the mood-optimised design will use a *less systematic (more heuristic) processing style* than users exposed to the original design, which means that they will *remember less of what they saw on the website.*

[94] As this calculation was not supported by the version of SPSS that was used, it was performed by use of University of Colorado's effect size calculator (Becker, 2015) (Appx. 12.9).

the effect of the design on DVs (and mood) is so small, the risk of *not* detecting the effect is more than 80% (a power level of less than 0.20) (cf. 8.7.1. and 8.8.3.).

As mentioned earlier, the more things feel wrong, the more systematic a processing style is used (Forgas, 1995; Schwarz, 2011). The results thereby suggest that the mood-optimised design made the subjects view the situation as more favorable and that they therefore did not pay as much attention to detail. They also suggest that the mood-optimised design improved the users' mood, since people in a good mood use a more heuristic processing style (ibid.). This will be further discussed in section 10.

Interestingly, the name and the price of the vase (memory question 1 and 3) were actually made larger and clearer in the mood-optimised design (in order to comply with the qualitative finding that a fluent design positively affects users' mood). While intuitively this could lead one to believe that this would make the treatment group remember the name and price of the vase better than the control group, the contrary effect is predicted by theory (ibid.) and was confirmed by the results of this study. The mood-optimised design lead the treatment group to remember less than the control group because they processed the information less systematically, thus, they did not go into as much depth with the processing. The control group was better at remembering what they had seen, even though the information was presented more clearly in the design given to the treatment group, because the control group had used a deeper processing style when reading the information.

It was expected that the effect of the design on processing style would be mediated by the design's effect on mood, which will be examined in the following section.

9.2.2. The Design's Effect on Mood

The design's effect on mood was tested to confirm the qualitatively derived hypothesis that the design improves users' mood and contribute to confirming that the relationships between design and the DVs are mediated by mood.

The tests found no effect for the design on mood (Appx. 12.7.2). Although all mood measuring instruments were included, the tests only found significant differences between the two experimental groups on the rating of *kumulere* (p=0.018). The rating was significantly lower for the treatment group (median=5, mean=5.14, SD=2.064) than for the control group (median=6, mean=5.86, SD=1.790), which pointed in the opposite direction of what was expected. However, since there was no significant difference between the experimental groups on the self-report of mood, which is considered the strongest mood-measuring instrument (cf. section 8.3.1.), or on any of the other non-self-report measures, this single

measure was not enough to conclude that the design actually deteriorated mood. Therefore, $H0_5$ could not be rejected.

9.3. Correlations

After testing whether the design had an effect on the DVs and mood, it was examined whether mood significantly accounts for variations in the DVs. This was done to confirm the theory-based hypotheses about a relationship between mood and judgments related to online shopping and to contribute to confirming that the relationships between the design and the DVs are mediated by mood. The results (Appx. 12.8) are presented in the following.

9.3.1. The Relationship between Mood and the Dependent Variables

The analysis suggested a positive relationship between mood and specific purchase intention (for the vase) and between mood and website design evaluation. No other significant relationships were found. In the following sections, the test results that led to the conclusions are presented.

9.3.1.1. Product Evaluation

No significant correlation was found between mood and product evaluation, although the relationship between the rating of *approbere* and website design evaluation was close to being significant (r=0.147, p=0.076). Therefore, $H0_2$ could not be rejected.

9.3.1.2. Specific Purchase Intention

The correlation analysis suggested a significant positive correlation between self-reported mood and specific purchase intention (r=0.197, p=0.017), between the rating of *approbere* and specific purchase intention (r=0.243, p=0.003) and between the ratings of *kumulere* and specific purchase intention (r=0.204, p=0.014). Moreover, the relationship between the rating of *endossere* and specific purchase intention was close to being significant (r=0.153, p=0.066). However, the correlations are weak (r>5) (Greasley, 2008), which confirms that the effect is difficult to detect with a small sample size. Since a relationship between mood and specific purchase intention was suggested by three (four if you include the rating of *endossere*) of the instruments, including the self-report instrument (the strongest mood measuring instrument (cf. 8.3.1.), $H0_4A$ could be rejected and $H1_4A$[95] confirmed.

[95] $H1_4A$ (as stated in section 8.6.): There is a *positive relationship* between mood and specific purchase intention.

This finding confirms previous findings (cf. section 4), which suggest that mood affects purchase intention (Donovan et al., 1994; Donovan & Rossiter, 1982; Park et al., 2005; Spies et al., 1997; Swinyard, 1993) and that people in a good mood are easier to persuade with weak arguments because they base their decision on heuristic cues (Bless, Bohner, Schwarz, & Strack, 1990; Bless, Mackie, & Schwarz, 1992; Forgas, 1995; Mackie & Worth, 1989; Schwarz, 2011).

9.3.1.3. General Purchase Intention

No significant correlation was found between mood and general purchase intention. Therefore, $H0_4B$ could not be rejected.

9.3.1.4. Website Design Evaluation

The correlation analysis suggested a significant positive correlation between self-reported mood and website design evaluation (r=0.202, p=0.014) and between the rating of *kumulere* and website design evaluation (r=0.182, p=0.028). Furthermore, the relationship between the rating of *endossere* and website design evaluation was close to being significant (r=0.141, p=0.090). Since the relationship was suggested by two (three if you include the rating of *endossere*) of the mood instruments, including the self-report, $H0_1$ could be rejected and $H1_1$[96] confirmed. However, the correlations were weak (r>5), which again confirms that the effect is difficult to detect with a small sample size

The above-mentioned result confirms previous findings (cf. section 4), which suggest that mood is positively correlated with people's evaluations (Forgas, 1995; Gorn, Goldberg, & Basu, 1993; Isen, Shalker, Clark, & Karp, 1978; Oatley, Keltner, & Jenkins, 2009; Schwarz & Clore, 1983). The findings can be explained by the Feelings-as-Information model by suggesting that users base their evaluation of the website design on their feelings at that moment of evaluation (Schwarz, 2011).

9.3.1.5. Processing Style

No relationship was found between mood and processing style. As mentioned previously, a negative relationship between mood and memory score was expected (suggesting that a good mood leads to a more heuristic processing style). A significant positive correlation was found between the rating of *kulance* and memory score (r=0.246, p=0.003), which pointed in the opposite direction of what was hypothesised.

[96] $H1_1$ (as stated in section 8.6.): There is a positive relationship between mood and website design evaluation.

Nevertheless, since this relationship was only suggested by one of the word-pleasantness ratings, and not by the self-report (the strongest), this finding was not enough to reject $H0_3$.

In the following section, the implications of the quantitative findings are discussed.

10. Discussion of the Quantitative results

The purpose of the quantitative study was to test whether the mood-optimised design raises the users' purchase intention (specific and general), product evaluation, website design evaluation and processing style, and that this effect is mediated by mood. The hypotheses were grounded in the results from the qualitative study (suggesting that the mood-optimised design would improve the users' mood) and results from other researchers (suggesting a positive correlation between mood and the DVs). The findings from the quantitative study confirm that the design leads to a more heuristic processing style and that mood is positively correlated with specific purchase intention and website design evaluation.

That no other relationships were found may be explained by the general finding of small effect sizes for the confirmed relationships (all effect sizes were between 0.141 and 0.243), which suggests that the design has a weak effect on mood and judgments. As explained previously, there is an 80% risk (a power level of 0.20) that an effect size of around 0.2 will not be detected in this study. The fact that only three of the 11[97] alternative hypotheses (27.27%) were confirmed fits with this prediction. Consequently, the hypotheses that were not confirmed cannot be rejected, since there is a risk that the hypothesised relationships exist, but were just not detected. They must therefore be tested on a larger sample.

The finding that the design has an effect on processing style and not any of the other variables can also be explained by the placement of all measurements *after* the shopping scenario instead of *during*, which may have caused validity issues (as discussed in section 8.8.). It is therefore possible that the mood-optimised design *did* alter the subjects' mood *during* exposure to the design, but that being introduced to the questionnaire afterwards neutralised or lowered this effect. Since the subjects were not asked to make any judgments until *after* the scenario, affect infusion would not have taken place until then, which can explain why there was no difference between the groups on the DVs (Forgas, 1995). Consequently, future studies should measure the DVs during the scenario and measure mood directly after the shopping scenario (before introduction to any other tasks).

[97] The number 11 includes the A and B versions of H_4 and H_9.

The implications of the various findings are discussed in detail in the following sections.

10.1. The Effect of the Design on Mood

The first question to be discussed is whether the qualitative insights led to the creation of a design that actually improves users' mood. Unfortunately, no relationship between design and mood was detected in this study. However, a positive relationship between design and processing style was found, and the positive relationship between mood and processing style, which has been found by numerous other researchers (Schwarz, 2011; Tiedens & Linton, 2001) suggests that the processing style measure can also be regarded as a valid mood measurement. If we accept this premise, it would be reasonable to conclude that the design did affect mood, but that this effect was not detected in the analysis of the data gathered by the mood-measuring instruments, either due to the low level of statistical power or due to the above-mentioned validity issues of the mood measurements.

If the hypothesis is tested on a larger sample with the measurements placed during the scenario instead of after, and $H0_5$ still cannot be rejected, it must be because the design simply does not have an effect on mood. This means that either the users were misunderstood in the qualitative study or the findings were just not integrated properly into the mood-optimised design due to lack of design skills or the limited access to photos that matched the qualitative findings.

10.2. The Relationship between Mood and Online Shopping Related Judgments

The next question is whether the theory-grounded predictions could be confirmed, namely that users' mood accounts for variations in their online shopping related judgments. This relationship was tested across the two experimental groups, and thus the results did not depend on the design's effect on mood. The results suggest that there is a relationship between mood and online shopping related judgments.

The results suggest that improving users' mood will raise their evaluation of the website. It may also raise their product evaluation (this was confirmed on a lower confidence level, $p=0.076$). This finding can be explained by the Feelings-as-Information model, which suggests that subjects base their evaluations on their feelings at that moment (Schwarz, 2011). Furthermore, a positive relationship between mood and specific purchase intention was identified. The finding of no significant correlation between mood and processing style puts a question mark to whether the effect of mood on specific purchase intention can be explained by Dual Process theory. As previously mentioned, the theoretical explanation behind the effect of mood on persuasion is based on the relationship between mood and processing style (since people in a positive mood use a more heuristic processing style, they are easier to persuade) (Bless, Bohner, Schwarz,

& Strack, 1990; Bless, Mackie, & Schwarz, 1992; Forgas, 1995; Mackie & Worth, 1989; Schwarz, 2011). However, due to the low level of statistical power, there may be a relationship between mood and processing style, which was just not detected. Only if the hypothesis is tested on a larger sample, and a positive relationship between mood and processing style still cannot be confirmed, it can be interpreted that the relationship between mood and specific purchase intention cannot be explained by Dual Process theory, but only by the Feelings-as-Information model.

The relationship between mood and purchase intention was only found to be present when it comes to buying the specific vase and not when it comes to buying any product on the website. This is surprising, since the majority of researchers who have confirmed a relationship between mood and purchase intention have used general, and not specific, purchase intention as a measure (Donovan & Rossiter, 1982; Swinyard, 1993). Yet, for discovering a relationship between mood and purchase intention on a website, Park et al. (2005) measured specific purchase intention (for a pair of Khaki pants). It is therefore likely that there is a positive relationship between mood and both kinds of purchase intentions. This suggestion is supported by the finding of a positive relationship between design and general purchase intention (discussed in section 10.3.).

If the hypothesis is tested on a larger sample, and $H0_4B$[98] still cannot be rejected, it may be because the users based their judgment on their general remembrance of products from Imerco (since Imerco is a well-known retailer) rather than what they saw in the scenario (where they only saw three vases). As previously mentioned (cf. section 4.2.), affect infusion is most likely to occur when people process new information, since people have direct access to existing judgments (stored in memory) about information they already know (Forgas, 1995). If the subjects' general purchase intention was based on stored judgments regarding Imerco's products and not new judgments about the products they had just seen, this explains why mood did not affect their judgment.

10.3. The Effect of the Design on Online Shopping Related Judgments

In the above two sections, the relationship between the design and mood and between mood and the DVs were discussed. In this section, the results that encompass both, namely the relationship between the design and the DVs, are discussed.

[98] $H0_4B$ (as stated in section 8.6.): There is *no relationship* between mood and general purchase intention.

The results suggest that there is a causal relationship between design and processing style, but not between design and the remaining DVs. The finding of an effect on processing style, along with the knowledge that the power level was only 0.2 and that the placement of the remaining instruments may have caused validity issues, suggests that the design may also affect the other DVs, but that this effect was not detected. Yet, the mean differences on product evaluation and specific purchase intention between the two groups point in the opposite direction of what was hypothesised. Even though these are far from being statistically significant (p=0.652 and 0.551), if these indications were confirmed they could be interpreted in light of the finding that people across both experimental groups generally did not like the vase (cf. section 9.1.). It is possible that that the design is not capable of changing people's evaluation of the vase because their baseline opinion of it is low.

Considering that a causal relationship was only confirmed between design and processing style, the rest of this discussion will focus on this finding. The results suggest that a large number of the subjects could not remember much of what they had seen in the online shopping scenario[99] despite having been informed to pay close attention to what they saw. This suggests that both groups used a low level of systematic processing and confirms the general finding that people are more prone to use a heuristic processing style (Forgas, 1995). Still, the results suggest that the treatment group used a more heuristic processing style, since they remembered significantly less. This confirmed the expectation that the mood-optimised design would put the users in a better mood than the original design and thereby inform them that the situation was more favorable, why less monitoring and processing effort was required (Forgas, 1995).

If the hypothesis is tested on a larger sample with the measurements placed *during* the scenario instead of after, and a causal relationship between design and mood is still not found, it must be concluded that the mood-optimised design stimulates a less systematic processing style regardless of the mood of the user. To identify what caused the change in processing style, if not mood, we must look at the cues included in the mood-optimised design. There are no theoretical reasons of why photos, background colour, advantages or a gift should alter the users' processing style if the effect is not mediated by a mood change. However, even though research on how metacognitive experiences (cf. section 7.3.1.) affect processing style is limited, Schwarz argues that it is likely that low/high processing fluency gives rise to a more systematic/heuristic processing style (Schwarz, 2011). He supports this suggestion by referring to a body of research that suggests that any information that conveys a situation as problematic/benign increases the likelihood that people engage in a systematic/heuristic processing style (Schwarz, 2004). In 2008, Song

[99] The average score for the memory questions was 2.42/4.

and Schwarz confirmed the relationship between fluency and processing style in a study where subjects were asked to answer the question "How many animals of each kind did Moses take on the Ark?". The group that read the question in an easy-to-read font rather than a hard-to-read font were significantly less likely to detect the error[100] (Song & Schwarz, 2008). Based on this finding, it can be interpreted that it may also be the higher fluency of the mood-optimised design that caused the change in processing style. However, even if this is the case, it appears as unlikely that the mood-optimised design did not cause any change in mood, since research has demonstrated that in general when things go smoothly we perceive the situation as more benign (Schwarz, 2011).

Since the results of both the quantitative and qualitative studies have now been presented and discussed, the general results are summarised in the next section, followed by a discussion of theoretical and managerial implications. The last section presents ethical reflections.

11. Conclusion and General Discussion

11.1. Summary of Results

The purpose of the present research was to explore how users' mood can be influenced by website design and to test whether improving their mood will increase their purchase intention, product evaluation and website design evaluation. To answer the research questions, the hardware dealer Imerco's webshop was used as a case. Firstly, a qualitative study was conducted to explore how various website elements influence mood. The findings from this study suggest that users' mood is positively affected by the following elements, which were named 'mood cues':

1. Calm, neat and unstaged photos accompanied by text and displaying the products in a nice, familiar, and decorative 'dream setting'.
2. A calm and light background colour, which creates a contrast between background and foreground.
3. Information that is fluent, which means it is easy to process because it is clearly presented (involves large size and high quality of text and photos, clear font and contrasts) and because only small amounts of information is presented at once.
4. A gift.
5. Clear information about service and product advantages.

[100] It was not Moses but Noah that brought animals to the Ark.

These insights were integrated into a 'mood-optimised' mock-up design (consisting of five web pages) and tested against the original design in a two-group experiment measuring purchase intention, product evaluation, website design evaluation and processing style. Processing style was measured because people in a good mood use a more heuristic (less systematic) processing style, which makes them easier to persuade with weak arguments.

The results suggest that the mood-optimised design makes users employ a more heuristic processing style, which indicates that the design also has an effect on mood. However, the statistical tests did not detect a difference between the experimental groups on the mood measurements or on the remaining dependent variables. It was found that mood significantly accounts for variations in website design evaluation and specific purchase intention (towards a specific product on the website), which confirms previous findings discovered in offline contexts. Nevertheless, no correlations were detected between mood and product evaluation, general purchase intention (towards all products on the site) and processing style.

Even though large parts of the predicted relationships were not confirmed, they cannot be rejected. This is because the effect size of the confirmed relationships (around 0.2) and the small sample size (146 subjects) suggest that the statistical power was around 0.20, which means there is around a 20% chance of detecting effects even though they do exist. Moreover, the placement of the mood measures *after* the shopping scenario instead of *during* may have caused validity issues, since the mood effect may have been neutralised or lowered before mood was measured. Consequently, further research should be carried out on a larger sample and the measurements should be placed during the scenario instead of after to confirm the effects that were predicted.

11.2. Theoretical Contributions and Limitations

The effect of mood on *online* judgments has not been thoroughly researched. The present study contributes to this branch of research by confirming a relationship between mood and online judgments. Furthermore, few studies have tested *how* mood can be affected positively online, and these have used confirmatory approaches to test the effect of *one* website element, e.g. photo size (Park et al., 2005). This means that there may be elements that affect mood online, which have not yet been discovered. Therefore, this study used a context sensitive approach by first exploring how various website elements influence mood. Then, after identifying five cues that affect mood, these were tested quantitatively. Due to resource constraints, all five mood cues were tested at once. Since the effect sizes of the cues may vary, future research should isolate each cue and test them separately.

Since this is a case study of Imerco's website, the findings are not necessarily generalisable to other research settings and there may be mood cues that were not identified. Consequently, future studies should test the robustness of present conclusions in other contexts and further explore how other website elements influence mood. Since time restrictions did not allow for exhaustive exploration of the meaning of each of the mood cues identified, future studies should also pursue this in order to understand the influence of variations of the cues (e.g. having people in the photos).

Future research should also test whether the effect of design on mood varies depending on the context, e.g. by testing the same mood cues on different types of websites. According to Sherman and Smith (1987) the likelihood of mood affecting offline purchase behavior is stronger in a specialty store, such as a clothing store, than in a store in which the costumer only makes routine purchases, e.g. purchasing razor blades (Sherman & Smith, 1987). This finding may also apply to online contexts. Even though the context of the present research was a specialty store, in the quantitative study the subjects were asked to imagine looking for a specific product (a vase), and this may have created a feeling of making a routine purchase. Therefore, if they had been asked to just browse through the website and then asked to rate their general purchase intention, a larger effect on mood might have taken place. The difference between general and specialty stores and between browsing and making routine purchases should therefore be further investigated.

Other researchers have found the effect of mood on judgment to be moderated by a number of variables. For example, it has been found that the effect of mood on judgment is more pronounced under conditions of reduced processing capacity induced by time pressure and competing task demands, because these conditions further increase reliance on heuristic cues (Mackie & Worth, 1989; Siemer & Reisenzein, 1998). Future studies should therefore test whether these moderators also apply in an online context and how these can be successfully applied for raising sales, e.g. by introducing time limited offers that urge the customers to make quick purchase decisions.

11.2. Managerial Implications

11.2.1. General Implications

The present research demonstrates that by applying psychological mood research to online marketing it is possible to gain valuable insights into how sales can be raised. The finding that the design of a website affects users' mood and thereby their online judgments has critical implications for online marketers, since it means that they can raise sales by designing their website to affect users' mood positively. Taking into account that the results of this study were obtained for Imerco's webshop, which constitutes a specific

context, similar webshops may also benefit from integrating the five mood cues that were identified. Still, the qualitative findings suggest that there are complex individual differences in users' reactions to the same website elements. The results suggest that some of these individual differences may depend on gender and culture. Furthermore, whether or not the website elements reflect the image of the brand influences how the users react on them. Therefore, managers should tailor their websites for their specific brand and users (based on in-depth knowledge about the users), and test the website's effect on the users' mood before publishing it. Another possible solution could be to use personalised designs for different types of users. Managers could do this by splitting the users into segments, based on the knowledge they have about them from cookies. In practice however, it would require access to a large amount of data and in-depth knowledge about how different types of website elements affect the various user segments.

11.2.2. Implications for Imerco

The identified mood cues can be directly implemented on Imerco's website to help raise sales and heighten the users' evaluations of the website and its products. Still, to confirm the practical value of the mood cues in a natural setting, it could be an advantage to first test each of them in a field experiment. This could be done by use of A/B split tests. This method includes using a tool that randomly exposes users to two variations of the same webpage, which makes it possible to compare the percentage of visitors who perform a specific action, e.g. buying or looking at a product (VWO, 2016). By using this method, actual behavior could be measured, which would make the conclusion more practicable. Participants would be recruited automatically when visiting the website, thus capturing the population of interest, and the experiment could continue until the sample is large enough to detect significant results. This type of experiment would increase the ecological validity, since it would make the findings more practically applicable to Imerco's website. Yet, the fact that the subjects would have no knowledge that they were in an experiment would require ethical reflections (Bryman, 2012).

Even though the results of a field experiment would be practically useful to Imerco, they would not be academically useful for answering the present research questions since it would not be possible to measure mood or any dependent variables besides from purchase behavior. There are disadvantages in terms of internal validity as well, since there are many uncontrollable factors in a field experiment that can create variance between the experimental groups (ibid.).

Now that limitations and implications have been discussed, this report is completed with some ethical reflections.

11.3. Ethical Reflections

The fact that some consumers have issues with controlling their own shopping behavior raises the question of whether it can be considered an ethical practice to aim at stimulating sales by manipulating consumer mood. On one hand, most consumers are fully aware that from the moment they walk into a store, the company's goal is to persuade them to spend money. This goal is often reflected in the sales clerk's positive and persuasive attitude and a store atmosphere that is designed to create a pleasant experience. On the other hand, since the online shopping experience is still new to many customers, it may be that they are less 'on guard' for manipulations when shopping online, which makes them more vulnerable. This raises the question of whether consumers should be informed if a website deliberately aims to manipulate their mood (e.g. via pop-ups). However, informing the consumers would ruin the purpose of trying to improve their mood, because their mood would not affect their judgment if they were aware of the source of it (cf. section 4.2.2.). This issue relates to a broader discussion of whether companies should take more social responsibility in their online advertising practices. Due to scope and space considerations of this research, the issue will solely serve as an interesting dilemma for future work.

12. Bibliography

Abu-Shanab, E. A. (2011). Education level as a technology adoption moderator. In *ICCRD2011 - 2011 3rd International Conference on Computer Research and Development* (Vol. 1, pp. 324–328). http://doi.org/10.1109/ICCRD.2011.5764029

AI-Therapy. (2016). Selecting the effect size. Retrieved February 29, 2016, from https://www.ai-therapy.com/psychology-statistics/sample-size-calculator

Andersen, B. H., & Hansen, E. J. (2000). *Et Sociologisk værktøj: Introduktion til den kvantitative metode.* København: Hans Reitzels forlag.

Baron, R. M., & Kenny, D. A. (1986). The moderator-mediator variable distinction in social psychological research: conceptual, strategic, and statistical considerations. *Ournal of Personality and Social Psychology*, *51*(6), 1173–1182. http://doi.org/10.1037/0022-3514.51.6.1173

Becker, L. (2015). Effect Size Calculators. Retrieved February 14, 2016, from http://www.uccs.edu/~lbecker/

Berger, E., & Mitchell, A. (1989). The Effect of Advertising on Attitude Accessibility, Attitude Confidence, and the Attitude-Behavior Relationship. *Journal of Consumer Research*, *16*(3). http://doi.org/http://dx.doi.org/10.1086/209213

Bless, H., Bohner, G., Schwarz, N., & Strack, F. (1990). Mood and Persuasion: A Cognitive Response Analysis. *Personality and Social Psychology Bulletin*, *16*(2), 331–345. http://doi.org/10.1177/0146167290162013

Bless, H., Mackie, D. M., & Schwarz, N. (1992). Mood effects on attitude judgments: independent effects of mood before and after message elaboration. *Journal of Personality and Social Psychology*, *63*(4), 585–595. http://doi.org/10.1037/0022-3514.63.4.585

Bower, G. H. (1981). Mood and memory. *The American Psychologist*, *36*(2), 129–148. http://doi.org/10.1037/0003-066X.36.2.129

Bryman, A. (2012). *Social Research Methods* (4th ed.). New York: Oxford University Press.

http://doi.org/10.4135/9781849209939

Buxton, B. (2007). *Sketching User Experiences*. San Francisco: Elsevier.

Cacioppo, J. T., & Petty, R. E. (1984). The elaboration likelihood model of persuasion. *Advances in Consumer Research*, *11*(1), 673–675. http://doi.org/10.1558/ijsll.v14i2.309

Cavanaugh, L. A. (2009). *Feeling good and doing better: How Specific Positive Emotions Influence Consumer Behavior and Well-being (Doctoral dissertation)*. Duke University.

Chaiken, S., & Ledgerwood, C. (2007). Dual process theories. In R. Baumeister & K. Vohs (Eds.), *Encyclopedia of social psychology.* CA: SAGE Publications.

Clore, G. L., Gasper, K., & Garvin, E. (2001). Affect as information. In J. P. Forgas (Ed.), *Handbook of Affect and Social Cognition*. Psychology Press.

Clore, G. L., & Parrott, W. (1994). Cognitive feelings and metacognitive judgments. *European Journal of Social Psychology*, *24*, 101–115. http://doi.org/10.1002/ejsp.2420240108

Craik, F. I. M., & Endel, T. (1975). Depth of Processing and the Retention of Words in Episodic Memory. *Journal of Experimental Psychology: General*, *104*(3). http://doi.org/10.1037/0096-3445.104.3.268

Cramer, D., & Howitt, D. (Eds.). (2004). Order Effect. In *The SAGE Dictionary of Statistics*. London: SAGE publications. Retrieved from http://srmo.sagepub.com.esc-web.lib.cbs.dk/view/the-sage-dictionary-of-statistics/SAGE.xml

Cyr, D., Head, M., Larios, H., & Pan, B. (2009). Exploring human images in website design: A Multi-Method Approach. *MIS Quarterly*, *33*.

Dansk, & Sprognævn. (1985). Svære ord og lette. *Nyt Fra Sprognævnet*, (1), 1–12.

David, M., & Sutton, C. (2007). *Social Research: The Basics*. SAGE Publications.

DeSteno, D., Li, Y., Dickens, L., & Lerner, J. S. (2014). Gratitude: A Tool for Reducing Economic Impatience. *Psychological Science*, *25*(6), 1262–1267.

http://doi.org/10.1177/0956797614529979

Donovan, R. J., & Rossiter, J. R. (1982). Store Atmosphere: An Environmental Psychology Approach. *Journal of Retailing*, *58*(1), 34. http://doi.org/Article

Donovan, R. J., Rossiter, J. R., Marcoolyn, G., & Nesdale, A. (1994). Store atmosphere and purchasing behavior. *Journal of Retailing*, *70*(3), 283–294. http://doi.org/10.1016/0022-4359(94)90037-X

Duque, M. J., Turla, C., & Evangelista, L. (2013). Effects of Emotional State on Decision Making Time. *Procedia - Social and Behavioral Sciences*, *97*, 137–146. http://doi.org/10.1016/j.sbspro.2013.10.214

Forgas, J. P. (1995). Mood and judgment: the affect infusion model (AIM). *Psychological Bulletin*, *117*(1), 39–66. http://doi.org/10.1037/0033-2909.117.1.39

Forgas, J. P. (2009). Feeling and Doing: Affective Influences on Interpersonal Behavior. *Psychological Inquiry: An International Journal for the Advancement of Psychological Theory*, *13*(1), 1–28. http://doi.org/10.1207/S15327965PLI1301_01

Gerrards-Hesse, A., Spies, K., & Hesse, F. W. (1994). Experimental inductions of emotional states and their effectiveness: A review. *British Journal of Psychology*. http://doi.org/10.1111/j.1750-3841.2009.01222.x

Gorn, G. J., Goldberg, M. E., & Basu, K. (1993). Mood, Awareness, and Product Evaluation. *Journal of Consumer Psychology*, *2*(3), 237–256. http://doi.org/http://dx.doi.org/10.1016/S1057-7408(08)80016-2

GraphPad. (2015). Sample size for nonparametric tests. Retrieved February 29, 2016, from http://www.graphpad.com/guides/prism/6/statistics/index.htm?stat_sample_size_for_nonparametric_.htm

Gratton, C., & Jones, I. (2014). *Research Methods for Sport Studies*. Ney York: Routledge.

Greasley, P. (2008). *Quantitative Data Analysis Using SPSS : An Introduction for Health and Social Sciences*. Open University Press.

Greenwald, M. K., Cook, E. W., & Lang, P. J. (1989). Affective judgment and psychophysiological response: Dimensional covariation in the evaluation of pictorial stimuli. *Journal of Psychophysiology*, *3*(1), 51–64.

Gross, J. J., & Levenson, R. W. (1995). Emotion elicitation using films. *Cognition & Emotion*, *9*(1), 87–108. http://doi.org/10.1080/02699939508408966

Göritz, A. S. (2007). The Induction of Mood via the WWW. *Motivation and Emotion*, *31*(1), 35–47. http://doi.org/10.1007/s11031-006-9047-4

Hansen, G. (2005). Experience and Emotion in Empirical Translation Research with Think-Aloud and Retrospection. *Meta: Journal Des Traducteurs*, *50*(2), 511. http://doi.org/10.7202/010997ar

Harwell, M. R. (2011). Research Design in Qualitative/ Quantitative/Mixed Methods. In C. F. Conrad & R. C. Serlin (Eds.), *The SAGE Handbook for Research in Education: Pursuing Ideas as the Keystone of Exemplary Inquiry*. SAGE Publications. http://doi.org/http://dx.doi.org.esc-web.lib.cbs.dk/10.4135/9781483351377

Hirshleifer, D., & Shumway, T. (2003). Hirshleifer and Shumway 2003 Good Day Sunshine: Stock Returns and the Weather. *The Journal of Finance*, *58*(3), 1009–1032. http://doi.org/10.1111/1540-6261.00556

Hume, D. (1896). Of the influencing motives of the will (part 3, section 3). In *A Treatise of Human Nature (Book 2)*. New York: Mineola.

IBM. (2010). How to calculate sample size for the Mann-Whitney-U test based on the sample size for the Student test. Retrieved from http://www-01.ibm.com/support/docview.wss?uid=swg21478271

Isen, A. M., & Geva, N. (1987). The influence of positive affect on acceptable level of risk: The person with a large canoe has a large worry. *Organizational Behavior and Human Decision Processes*, *39*(2), 145–154. http://doi.org/10.1016/0749-5978(87)90034-3

Isen, A. M., Johnson, M. M., Mertz, E., & Robinson, G. F. (1985). The influence of positive affect

on the unusualness of word associations. *Journal of Personality and Social Psychology*, *48*(6), 1413–1426. http://doi.org/10.1037/0022-3514.48.6.1413

Isen, A. M., & Means, B. (1983). The influence of positive affect on decision-making strategy. *Social Cognition*, *2*(1), 18–31. http://doi.org/10.1521/soco.1983.2.1.18

Isen, A. M., Shalker, T., Clark, M., & Karp, L. (1978). Affect, accessibility of material in memory, and behavior: a cognitive loop? *Journal of Personality and Social Psychology*, *36*(1), 1–12. http://doi.org/10.1037/0022-3514.36.1.1

Isen, A. M., & Shalker, T. E. (1982). The Effect of Feeling State on Evaluation of Positive, Neutral, and Negative Stimuli: When You "Accentuate the Positive," Do You "Eliminate the Negative"? *Social Psychology Quarterly*, *45*(1), 58. http://doi.org/10.2307/3033676

Johnson, A. R., & Stewart, D. W. (2010). A reappraisal of the role of emotion in consumer behavior: Traditional and contemporary approaches. *Review of Marketing*, *1*, 3–34. http://doi.org/10.1108/S1548-6435(2004)0000001005

Johnson-Laird, P. N. (2010). Mental models and human reasoning. *Proceedings of the National Academy of Sciences of the United States of America*, *107*(43), 18243–50. http://doi.org/10.1073/pnas.1012933107

Jones, N. a., Ross, H., Lynam, T., Perez, P., & Leitch, A. (2011). Mental Model an Interdisciplinary Synthesis of Theory and Methods. *Ecology and Society*, *16*(1), 46–46. http://doi.org/46

JQuery. (2015). Accordion. Retrieved February 28, 2016, from https://jqueryui.com/accordion/

Kahneman, D. (2002). Maps of bounded rationality: a perspective on intuitive judgment and choice. *The Sveriges Riksbank Prize in Economic Sciences in Memory of Alfred Nobel*, (December), 449–489. http://doi.org/10.1037/0003-066X.58.9.697

Kearney, a. R., & Kaplan, S. (1997). Toward a Methodology for the Measurement of Knowledge Structures of Ordinary People: The Conceptual Content Cognitive Map (3CM). *Environment and Behavior*, *29*, 579–617. http://doi.org/10.1177/0013916597295001

Kolb, B. (2008). *Marketing Research*. London, England: SAGE Publications.

http://doi.org/http://dx.doi.org.esc-web.lib.cbs.dk/10.4135/9780857028013

Larsen, R. J., & Sinnett, L. M. (1991). Meta-Analysis of Experimental Manipulations: Some Factors Affecting the Velten Mood Induction Procedure. *Personality and Social Psychology Bulletin*, *17*, 323–334. http://doi.org/10.1177/0146167291173013

Laudon, K. C., & Traver, C. G. (2015). The online retail sector. In *E-commerce: Business, technology, society* (11th ed.). Boston: Pearson.

Lerner, J. S., Han, S., & Keltner, D. (2007). Feelings and Consumer Decision Making: Extending the Appraisal-Tendency Framework. *Journal of Consumer Psychology*, *17*(3), 181–187. http://doi.org/10.1016/S1057-7408(07)70027-X

Lerner, J. S., & Keltner, D. (2000). Beyond valence: Toward a model of emotion-specific influences on judgement and choice. *Cognition & Emotion*, *14*(4), 473–493. http://doi.org/10.1080/026999300402763

Lerner, J. S., Li, Y., & Weber, E. U. (2013). The financial costs of sadness. *Psychological Science*, *24*(1), 72–79. http://doi.org/10.1177/0956797612450302

Lerner, J. S., Small, D. a., & Loewenstein, G. (2004). Heart strings and purse strings: Carryover effects of emotions on economic decisions. *Psychological Science*, *15*(5), 337–341. http://doi.org/10.1111/j.0956-7976.2004.00679.x

Lerner, J. S., & Tiedens, L. Z. (2006). Portrait of the angry decision maker: How appraisal tendencies shape Anger's influence on cognition. *Journal of Behavioral Decision Making*. http://doi.org/10.1002/bdm.515

Mackie, D. M., & Worth, L. T. (1989). Processing deficits and the mediation of positive affect in persuasion. *Journal of Personality & Social Psychology*, *57*(1), 27–40. http://doi.org/10.1037/0022-3514.57.1.27

Miles, M., & Huberman, M. (1994). *Qualitative data analysis: An expanded sourcebook* (2nd ed., Vol. 20). USA: SAGE Publications.

Nielsen, J. (2011). Why You Only Need to Test with 5 Users. Retrieved February 8, 2016, from

http://www.nngroup.com/articles/why-you-only-need-to-test-with-5-users/

Oatley, K., Keltner, D., & Jenkins, J. M. (2006). *Understanding emotions* (2nd ed.). Wiley-Blackwell.

Olsen, H. (1998). Spørgsmålet som Tekst. In *Tallenes talende tavshed*. Copenhagen: Akademisk Forlag.

Park, J., Lennon, S. J., & Stoel, L. (2005). On-line product presentation: Effects on mood, perceived risk, and purchase intention. *Psychology and Marketing*, *22*(9), 695–719. http://doi.org/10.1002/mar.20080

Putrevu, S. (2004). Communicating with the Sexes : Male and Female Responses to Print Advertisements. *Journal of Advertising*, *33*(3), 51–62. http://doi.org/10.1080/00913367.2004.10639168

RANDOM.ORG. (2016a). List Randomizer. Retrieved February 29, 2016, from https://www.random.org/lists/

RANDOM.ORG. (2016b). True Random Number Service. Retrieved February 9, 2016, from https://www.random.org/

Reber, R., & Schwarz, N. (1999). Effects of perceptual fluency on judgments of truth. *Consciousness and Cognition*, *8*(3), 338–342. http://doi.org/10.1111/1467-9280.00008

Saegert, J., & Young, R. (1981). Comparison of Effects of Repetition and Levels of Processing in Memory For Advertisements. *Advances in Consumer Research*, *9*.

Sartre, J.-P. (2004). *Sketch for a Theory of the Emotions*. (P. Mairet & M. Warnock, Eds.). London and New York: Routledge.

Saunders, M., Lewis, P., & Thornhill, A. (2009). *Research Methods for Business Students*. *Research methods for business students* (5th ed.). Pearson.

Schwartz, A. (2009). Dual-process theory. In M. Kattan (Ed.), *Encyclopedia of medical decision making.* CA: SAGE Publications.

Schwarz, N. (2004). Metacognitive Experiences in Consumer Judgment and Decision Making. *Journal of Consumer Psychology*, *14*(4), 332–348. http://doi.org/10.1207/s15327663jcp1404_2

Schwarz, N. (2011). Feelings-as-Information theory. In P. A. M. Van Lange, A. W. Kruglanski, & E. T. Higgins (Eds.), *Handbook of Theories of Social Psychology* (Vol. 1, pp. 289–308). London: SAGE Publications. http://doi.org/10.1007/978-1-4614-0158-2

Schwarz, N., & Clore, G. L. (1983). Mood, misattribution, and judgments of well-being: Informative and directive functions of affective states. *Journal of Personality and Social Psychology*, *45*(3), 513–523. http://doi.org/10.1037/0022-3514.45.3.513

Schwarz, N., Sanna, L. J., Skurnik, I., & Yoon, C. (2007). Metacognitive Experiences and the Intricacies of Setting People Straight: Implications for Debiasing and Public Information Campaigns. *Advances in Experimental Social Psychology*, *39*(06), 127–161. http://doi.org/10.1016/S0065-2601(06)39003-X

Sherman, E., & Smith, R. B. R. (1987). Mood states of shoppers and store image: promising interactions and possible behavioral effects. *Advances in Consumer Research*, *14*(1), 251–254.

Siemer, M., & Reisenzein, R. (1998). Effects of Mood on Evaluative Judgements: Influence of Reduced Processing Capacity and Mood Salience. *Cognition & Emotion*, *12*(6), 783–805. http://doi.org/10.1080/026999398379439

Simon, S. (2000). Stats: Sample size for Mann-Whitney U. Retrieved February 29, 2016, from http://www.pmean.com/00/mann.html

Song, H., & Schwarz, N. (2008). Fluency and the Detection of Misleading Questions: Low Processing Fluency Attenuates the Moses Illusion. *Social Cognition*, *26*(6), 791–799. http://doi.org/10.1521/soco.2008.26.6.791

Spies, K., Hesse, F., & Loesch, K. (1997). Store atmosphere, mood and purchasing behavior. *International Journal of Research in Marketing*, *14*(1), 1–17. http://doi.org/10.1016/S0167-8116(96)00015-8

Swinyard, R. (1993). The Effects of Mood, Involvement, and Quality of Store Experience on Shopping Intentions. *Journal of Consumer Research*, *20*(2), 271–280. http://doi.org/10.1086/209348

Tiedens, L. Z., & Linton, S. (2001). Judgment under emotional certainty and uncertainty: the effects of specific emotions on information processing. *Journal of Personality and Social Psychology*, *81*(6), 973–988. http://doi.org/10.1037/0022-3514.81.6.973

Trochim, W. M. K., & Donnelly, J. P. (2008). *The Research Methods Knowledge Base*. *WebRef* (3rd ed.). Mason, Ohio: Cengage Learning.

Tversky, A., & Kahneman, D. (1973). Availability: A heuristic for judging frequency and probability. *Cognitive Psychology*, *5*(2), 207–232. http://doi.org/10.1016/0010-0285(73)90033-9

Velten, E. (1968). A laboratory task for induction of mood states. *Behaviour Research and Therapy*, *6*(4), 473–482. http://doi.org/10.1016/0005-7967(68)90028-4

Verheyen, C., & Goritz, A. S. (2009). Plain texts as an online mood-induction procedure. *Social Psychology*, *40*(1), 6–15. http://doi.org/10.1027/1864-9335.40.1.6

VWO. (2016). The Complete Guide to A/B Testing. Retrieved March 13, 2016, from https://vwo.com/ab-testing/

Westermann, R., Stahl, G., & Hesse, F. W. (1996). Relative effectiveness and validity of mood induction procedures: a meta- analysis. *European Journal of Social Psychology*, *26*(4), 557–580. http://doi.org/10.1002/(SICI)1099-0992(199607)26:4<557::AID-EJSP769>3.0.CO;2-4

Winkielman, P., & Zajonc & Norbert Schwarz, R. B. (1997). Subliminal Affective Priming Resists Attributional Interventions. *Cognition & Emotion*, *11*(4), 433–465. http://doi.org/10.1080/026999397379872

Yoon, C. (1997). Age Differences in Consumers' Processing Strategies: An Investigation of Moderating Influences. *Journal of Consumer Research*, *24*(3), 329–342. http://doi.org/10.1086/209514