Master thesis

First Impression Website Attractiveness: a Consumer Neuroscience Approach

Leander Rotshuizen
S1759590

Master's programme

Master Business Administration

Strategic Marketing and Business Information

Supervisors
Dr. Van der Lubbe
Dr. Constantinides

Date

21-01-2018

ABSTRACT

This study examines whether it is possible to make well-founded first impression attractiveness judgements of websites in only 50ms by comparing it to a 500ms exposure duration. Visual complexity (VC) and prototypicality (PT) are important factors when making first impression attractiveness judgements according to Leder et al. (2004) and are therefore included in this study. This study examined besides the effect of VC and PT on attractiveness also the effect of both factors on visual attention, as visual attention is expected to be related to the choice of a website. The experiment displayed two webpages (with different VC and PT properties) at a time and the participants had to indicate which website they found the most attractive. Presentation time of the websites was varied (500ms vs 50ms) in order to compare the effect of VC and PT at both durations. Moreover, electroencephalogram (EEG) and electrooculogram (EOG) data was collected in order to measure visual attention and eye movements. Results showed that there is a negative effect of high VC and positive effect of high PT on first impression visual appeal judgements of websites in both the 500ms and 50ms condition. The effect of high PT is positive in the behavioural data but negative in the EEG data, suggesting that atypical websites draw visual attention but are not perceived as visually appealing. Finally, results suggest that the effect of VC is as fast and strong as PT in both exposure durations. The findings in this paper are in particular important for web designers to take into account when developing or optimizing web designs, as adaptation of the findings in this study will increase first impression attractiveness judgements of their websites.

1. Introduction

This paper has its roots in human-computer interaction (HCI) research and is interested in the first impression attractiveness[1] that is generated by individuals when perceiving websites. In psychology, first impressions are described as the mental image of a person that is generated in the event of meeting another person for the first time (Nevid, 2012). This definition can be extended to not only include the mental images of other people but also the mental image of websites (Lindgaard, Fernandes, Dudek, & Brown, 2006, 2011; Tuch, Presslaber, Stöcklin, Opwis & Bargas-Avila, 2012). The mental images of persons can be generated extremely quickly. As an example, Willis and Todorov (2006) showed that people can make trait inferences of individuals (attractiveness, likeability, trustworthiness, competence, and aggressiveness) by examining their facial appearances even within an exposure duration of 100ms. Bar, Neta and Linz (2006) conducted a similar experiment which revealed that first impressions of personality traits (threatening and nonthreatening) generated in 39ms were similar to the first impressions generated in a 1700ms exposure. In other words, these researchers revealed that first impressions based on facial appearances are generated within the first 100ms.

1.1 Website first impression visual appeal studies

Similar results as described above seem to also apply for the mental images of websites. The topic of perceived attractiveness of websites that are created in short exposure durations has been researched before (Lindgaard et al., 2006, 2011; Tractinsky, Cokhavi, Kirschenbaum, & Sharfi, 2006; Tuch et al., 2012). Lindgaard et al. (2006) are the pioneers in first impression visual appeal research related to websites. The study that these researchers conducted consists out of three similar experiments. In experiment one and two participants viewed and rated multiple homepages of websites at an exposure duration of 500ms, each homepage was viewed and rated twice. This exposure duration was intended to: (1) be long enough to generate a first impression based on visual appeal. (2) be not sufficient amount of time that the content of the homepage could be rated. Their first experiment consisted of 100 homepages and experiment two consisted of the 25 most attractive and 25 most unattractive

[1] Attractiveness, visual appeal and aesthetics are used interchangeably in this paper.

homepages of experiment one. After each homepage exposure the respondents rated the homepage from very unattractive to very attractive on a 0-100 rating scale. The purpose of these experiments was to assess the consistency of the visual appeal ratings in these short exposure durations and to explore visual appeal characteristics. The results in the first and second experiment indicated that the correlation between the mean first rating and mean second rating of a website correlated highly with a squared Pearson correlation coefficient of 0.97. These results indicate that the visual appeal ratings were stable at an exposure duration of 500ms. This suggests that individuals are able to get reliable first impressions of websites within 500ms. This result is similar in other studies that replicated or extended the study from lindgaard et al. (2006) (e.g. Van Schaik & Ling, 2009; Tractinsky et al., 2006)

In the third experiment, Lindgaard et al. (2006) conducted a between-subject comparison based on two different exposure durations: a short exposure duration of 50ms and a long exposure duration of 500ms. The findings indicated that the ratings of both the 50ms and the 500ms condition were consistent between mean first and mean second rating. This suggests that individuals are able to get reliable first impressions of websites even within 50ms. The main difference between the two conditions was the higher variability between subjects in the 50ms condition compared to the 500ms condition. According to Lindgaard et al. (2006) this is possibly due to the fact that in the 500ms condition the participants were able to interpret more content of the website. Other researchers replicated or extended this research and found similar results regarding the 50ms condition (e.g. Lindgaard et al., 2008, 2011; Tuch et al., 2012). These results are interesting but the majority of the studies in this field did not use an appropriate masking method to limit the time that the presented homepages are perceived. The consequence is that the set of website stimuli are perceived for a longer duration than the intended 50ms. More information regarding masking is given in section 1.4.

1.2 The halo effect

The previous mentioned results are not only interesting but also very important. One of the most notable forms of first impressions is visual appearance according to Tractinsky et al., (2006). The subjective attractiveness based on this visual appearance is created at first and is critical for generating a more thorough opinion and influences subsequent experiences and can therefore be considered as a 'halo effect' (Lindgaard et al., 2006, 2011; Lorenzo-Romero et al., 2013; Van Schaik & Ling, 2009). The halo effect can be defined as 'a rater's failure to discriminate among conceptually distinct and potentially independent attributes, with the result that individual attribute ratings co-vary more than they otherwise would' (Leuthesser, Kohli & Harich, 1995, pp. 58). This means that the first impression of the website attractiveness will also influence other attributes of the website, for instance usability and credibility (Lindgaard et al., 2011; Robins & Holmes, 2008). An individual visiting a website will first create a first impression based on the attractiveness of the website. If this first impression is negative it may result in the individual leaving the website or as the halo effect suggests, can lead to a negative impression of other aspects of the website and/or organisation. A positive first impression on the other hand can ultimately lead to the person becoming a regular visitor or customer and evaluating other aspects of the website and/or organisation higher.

1.3 Aesthetic perception processing

It is important how aesthetic or attractiveness judgements are formed by users of websites, as this knowledge will help improve website attractiveness. Leder et al. (2004) developed a theoretical framework that deals with the process of aesthetic (art) stimuli. This information-processing stage model of aesthetic processing is shown in figure 1 on the next page. The model includes five different processing stages that occur sequential: (1) perceptual analysis, (2) implicit information integration, (3) explicit classification, (4) cognitive mastering and finally (5) evaluation. The first two processing stages are related to the more basic cognitive processes and intuition. These processing stages do not require conscious evaluation but are more related to a 'gut feeling' as mentioned by Tuch et al. (2012). In the first stage, perceptual analysis, the stimulus is analysed based on perceptual features of a stimulus (e.g. colour and complexity). The second stage, implicit information integration, is related to previous history and experiences of the user. So familiarity and prototypicality shape the perception of attractiveness. Stage three and four are related to the higher cognitive processes (e.g. expertise and knowledge) which are not relevant in this study. The final stage is evaluation where the processed stimulus is evaluated.

The process of aesthetic perception processing requires a sufficient amount of time to achieve the full processing as depicted in the information-processing stage model of aesthetic processing. Time constraints limit the aesthetic perception processing and therefore result in information being processed in only the lower stages of the model. As this study is interested in first impressions where stimuli are shown for a fraction of a second, the perceptual analysis stage and implicit information integration stage are expected to be important. Therefore the effects of visual complexity (VC) from the perceptual analysis stage and prototypicality (PT) from the implicit information stage on attractiveness are examined. Both factors were introduced into the field of first impression website attractiveness by Tuch et al.

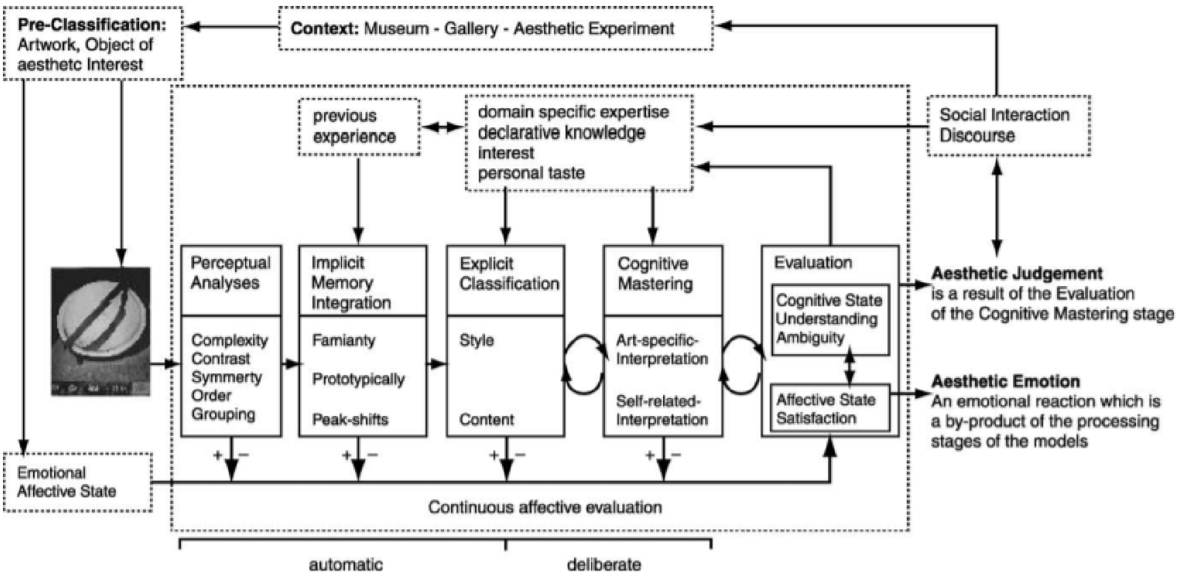


Figure 1 The information-processing stage model of aesthetic processing (Leder et al., 2004).

1.3.1 Visual Complexity

Visual Complexity (VC) plays an important role in visual appeal judgments according to Tuch et al. (2009, 2011, 2012). Tuch et al. (2012) and Xing and Manning (2005) state that the term VC is difficult to define. In line with Tuch et al. (2012) this study is also not interested in creating an objective definition of VC, the importance lies mainly in the subjective perceived complexity of the respondents. An example of a complex website is a website with a lot of elements (e.g. text, links and images), see appendix 5.1 for an example. The results of Tuch et al. (2012) show that first impressions of websites with low VC were rated higher than websites with a high amount of VC. This effect of VC was even apparent at an exposure duration of 17ms. Thielsch and Hirschfeld (2012) found similar results when testing the effect of low-spatial frequencies on aesthetic judgements of websites at a short exposure duration. Low-spatial frequencies can be regarded as VC as this conveys only the global information of a stimulus. In contrast, high-spatial frequencies encode the fine details of stimulus and therefore take much longer to process. VC is a perception and is included in the perceptual analysis which is the first processing stage that is involved with making attractiveness judgements according to Leder et al. (2004). Therefore it is expected that the effect of VC on visual appeal judgements of websites will be fast. Concluding, VC is a perception and stimuli with a low amount of VC are easier to process as they contain less information which needs to be processed, this may lead to a more positive judgement (Reber et al., 2004).

1.3.2 Prototypicality

Prototypicality (PT) is defined by Leder et al. (2004) as: 'the amount to which an object is representative of a class of objects' (pp. 496). PT is represented by mental models that are built through experience. Therefore the familiarity with certain objects and stimuli can explain its PT. Due to the extensive interaction of users with the internet, certain expectations are created of how a website should look. The study of Roth, Schmutz, Pauwels, Bargas-Avila and Opwis (2010) revealed that internet users have several distinct models for a variety of webpages (e.g. company websites and webshop). This means that web users have certain expectations related to the position of web objects (e.g. search field and navigation). According to the researchers the users agree about the location of many web objects, but not all. It is for instance agreed upon that the search bar of company websites is located on the top right and the navigation menu on the left side of the homepage (Roth, Schmutz, Pauwels, Bargas-Avila & Opwis, 2010). An example of a high and low PT website is given in appendix 5.1. Tuch et al. (2012) found that high PT was positively related with aesthetic perception. This effect of high PT was even apparent at an exposure duration of 17ms.

1.4 Afterimages

Even though there are numerous researchers that assessed if first impressions of websites can be made within 50ms, it is still important to conduct a similar study for two important reasons. The first being that doubts may be raised regarding the control of afterimages in the majority of studies in this field (e.g. Lindgaard et al., 2006, 2008, 2011). Due to a retinal afterimage and the afterimage on the computer monitor, the perception of a stimulus remains for approximately 250ms after the stimulus has disappeared (Goldstein, 2009). So the consequence is that in the majority of studies in this field the set of website stimuli are perceived for a longer duration (300ms) than intended (50ms). As the website stimuli were perceived for a longer duration than 50ms, it is logical that it is easier to judge the websites on attractiveness compared to when the websites were truly perceived for 50ms. The effect of an afterimage can be countered by following a stimulus with a mask (Breitmeyer & Öğmen, 2006; Enns, & Di Lollo, 2000). Masking can be defined as 'the reduction of the visibility of one stimulus, called the target, by a spatiotemporally overlapping or contiguous second stimulus, called the mask' (Breitmeyer & Öğmen, 2006, pp. 2). This study applied a backward mask to control for afterimages. In backward masking, the stimuli onset are followed by the mask's onset. With the introduction of a backward mask in this study we re-examined the results of Lindgaard et al. (2006), amongst others, that 50ms is enough to make a first impression attractiveness judgement of a website.

1.5 Visual attention

Another important reason for conducting this study was that the current studies in this field are solely focused on the concept attractiveness, even though visual attention can provide meaningful results. This is especially true as according to Pieters and Warlop (1999) it is believed by marketers and academics that the visual attention of consumers is related with the final brand choice of a product. This is for instance the reason that expressive product packaging is used by many brands. It is expected that this result is similar for websites, i.e. websites that draw a lot of visual attention are chosen over websites that do not draw visual attention. It is not entirely clear how this visual attention relates to attractiveness. Current research in this field is mainly focused on the attractiveness and visual attention of individuals faces. For instance, Seidman and Millar (2013) found that attractive individuals on Facebook profiles receive more attention in comparison with unattractive individuals. Maner et al. (2003) found similar results, as more attention was given to attractive opposite-gender faces in comparison with unattractive opposite-gender faces. There is however also proof that objects or persons that draw a lot of visual attention are not necessarily perceived as attractive. Van Hoof, Crawford and Van Vugt (2010) have shown for instance that both attractive and unattractive faces of the opposite sex receive more attention in comparison with average faces. The researchers suggest that their measure of attention was therefore more driven by physical distinctiveness rather than by attractiveness. Another example is that threatening stimuli, of for instance spiders and snakes, are able to capture our attention faster than non-threatening stimuli (Öhman, Flykt & Esteves, 2001). Above results show that attractive individuals receive more visual attention in comparison with unattractive individuals. There are however exceptions and it is unknown if this relation between attractiveness and visual attention is also apparent when websites are perceived instead of individuals faces. Therefore we are interested in which kind of websites draw visual attention (regarding VC and PT) and whether a discrepancy exist between visual attention and attractiveness when perceiving websites. An electroencephalogram (EEG) was applied in this research as the measured PO7 (left parietal occipital) and PO8 (right parietal occipital) electrodes can be used to indicate visual attention (e.g. Foxe, Simpson & Ahlfors, 1998; Thut, Nietzel, Brandt, & Pascual-Leone, 2006). Together with this EEG data, electrooculogram (EOG) data was gathered. The EOG's main applications was measuring eye movement and correcting for eye blinks in the EEG. In short, the EEG and EOG will assess whether websites that draw a lot of visual attention are also perceived as attractive.

The EEG data was analysed by computing and analysing event-related potentials (ERPs). ERPs can be defined as the electrical potentials that are recorded at an individual's scalp and are related to an internal or external event (Celesia & Brigell, 1992; Luck, 2012; Van der Lubbe et al., 2016). These events can range from interpreting stimuli, giving a response, making decisions, etc. The approach of calculating the ERPs is by averaging all time-locked activity that is relevant to a certain event and category (Van der Lubbe et al., 2016). For instance, the raw EEG data that appears when a participant indicates his preference for a website (event) that has high VC and high PT (category) is averaged.

When the data is averaged, the relevant data remains as it is time-locked to the event and noise is removed as it is only temporal (Van der Lubbe et al., 2016). In previous example, the ERP of only high VC and high PT websites remain. ERPs can be transformed into ERLs by calculating the contra-ipsilated hemispherical difference for the relevant electrodes (Van der Lubbe & Utzerath, 2013). In ERLs the left side of the brain can be compared with the right side of the brain as certain tasks are expected to have a larger effect on either the left or right side of the hemisphere. The focus of the EEG analysis was on the N2pc or posterior contralateral negativity (PCN). The PCN is an ERL measure that can be used to indicate visual attention (Eimer, 1996; Van der Lubbe, Jaśkowski, Wauschkuhn, & Verleger, 2001).

Concluding, this research is interested in answering three main questions. The first question will replicate the research of Tuch et al. (2012) and asks how high VC and high PT influence first impression visual appeal judgements of websites. The second question is how this influence of high VC and high PT relates to the exposure duration of a website. In other words, are the effects of high VC and high PT stronger or weaker in a 500ms condition compared to a 50ms condition? Note that above questions will also assess whether well-founded first impression attractiveness judgements can be made within an exposure duration of 50m. This is due to the fact that there will only be an effect of high VC and/or high PT in the 50ms condition when it is possible to make attractiveness judgements in 50ms. The judgements are acknowledged as well-founded when the judgements are not based on chance. This is important as it is also possible to make a judgement of a stimulus that is not actually perceived. The final question is whether there is a discrepancy between attractiveness and visual attention when judging websites on attractiveness that differ in the degree of VC and PT.

1.6 Hypothesis and predictions

VC was expected to have a negative effect on visual appeal judgements as stimuli with a low amount of VC are easier to process as they contain less information which needs to be processed. This may lead to a more positive judgement according to the fluency effect (Reber et al., 2004). The fluency effect proposes that the more fluent a stimulus can be processed, the more attractive they are perceived. Tuch et al. (2012) found that websites with low VC were indeed rated higher than websites with a high amount of VC. This effect of VC was even apparent at an exposure duration of 17ms. Thielsch and Hirschfeld (2012) found similar results when testing the effect of low-spatial frequencies (website filtered to global layout) on aesthetic judgements of websites at a short exposure duration. This resulted in the first hypothesis:

H1: Visual complexity has a negative effect on attractiveness judgements of websites even within 50ms.

PT on the on the hand was expected to have a positive effect on visual appeal judgements as Winkielman, Halberstadt, Fazendeiro and Catty (2006) have found that prototypical stimuli are easier to process than atypical stimuli. According to the fluency effect of Reber et al. (2004) this would result in higher perceived attractiveness. Moreover, it is shown by many researchers that prototypical objects or stimuli are preferred over non-prototypical objects or stimuli (Reber et al., 2004). Regarding website visual appeal judgements, Tuch et al. (2012) found that PT was positively related with aesthetic perception. This resulted in the second hypothesis:

H2: High prototypicality has a positive effect on attractiveness judgements of websites even within 50ms

It was expected that the effect of high VC and high PT is stronger in the 500ms condition when compared to the 50ms condition. A shorter exposure duration will increase the uncertainty of the respondent regarding their judgements and chance is therefore expected to play a larger role. This resulted in the third hypothesis:

H3: The effect of high visual complexity and high prototypicality on attractiveness judgements of websites is stronger in the 500ms condition in comparison to the 50ms condition

As mentioned, many researchers have claimed that individuals can make well-founded first impression visual appeal judgements of websites within 50ms and that these judgements relate with longer exposures. However, replication was required because the majority of studies in this field did not apply an appropriate mask. The expectations above indicate that it was still expected in this study that individuals can make well-founded first impression visual appeal judgements of websites within 50ms when a mask is applied. This is due to the fact that emotion research has already proven that first impressions of persons can be made within 50ms. For instance, the research from Bar et al. (2006) revealed that first impressions of personality traits are already generated within 39ms. Therefore it is expected that well-founded first impressions visual appeal judgements can be made within an exposure duration of 50ms. Hypothesis one, two and three could assess this expectation, as there will only be an effect of high VC and/or high PT when it is possible to make attractiveness judgements within 50ms.

VC is a perception and is included in the perceptual analysis which is the first processing stage that is involved with making attractiveness judgements according to Leder et al. (2004). Therefore it was expected that the effect of VC on visual appeal judgements of websites will be fast. PT is included in the implicit memory integration stage, which is the second processing stage when making aesthetic judgements according to Leder et al. (2004). According to the researchers, this stage requires implicit memory effects. The memory effects are implicit due to the fact that in this stage the individual making attractiveness judgements does not have to become aware in order to influence their judgements, the effect of memory is automatic. Because PT requires memory to some extent it was expected that the effect of PT on attractiveness judgements will be slower than VC. Tuch et al. (2012) showed that with an exposure duration of 17ms the effect of PT was less pronounced than VC. With an increasing exposure duration the effect of PT also gained strength. Similar results were expected in this study and therefore the fourth hypothesis was tested:

H4: The effect of high visual complexity on attractiveness judgements of websites is faster than the effect of high prototypicality

Finally, it is expected that websites that are perceived as attractive also receive more visual attention. This means that high VC has a negative effect and that high PT has a positive effect on visual attention when perceiving websites. This is due to the fact that emotion research has indicated that attractive individuals/faces receive more visual attention than unattractive or average individuals/faces (Maner et al., 2003; Seidman and Millar, 2013; Van Hoof et al., 2010). Similar results were expected when websites are perceived instead of people. This resulted in the fifth hypothesis:

H5: There is no discrepancy between attractiveness judgements of websites and visual attention for high/low VC and/or PT websites

2. Methodology

2.1 Participants

This study involved 17 participants with normal vision or corrected-to-normal vision, no colour blindness and no neurological diseases. The behavioural data of all 17 participants was used and the PCN of 15 participants was used after exclusion of two participants due to noisy EEG recordings. Participants were required to fill in a participant registration form that included, amongst others, information about the demographics, visibility, medication and neurological and psychiatric history of the participants. To assess the colour-blindness of the participants, the colour-blindness test of Ishihara (1976) was used. Participants were asked to report numbers that are depicted in multi-coloured pictures. Participants were granted a pass when they reported all numbers as correct. The acuity of the participants eyes were tested with the Freiburg Visual Acuity and Contrast Test of Bach (2006). In this test participants were standing 3 meters in front of a computer screen where 18 Landolt rings would be displayed. The Landolt rings ranged in size dependent on how many correct answers the participant has given. Curtains were closed and lights were partially turned off in order to exclude artefacts and keeping the conditions similar for every participant. The results of the acuity test showed that two respondents had a weak right eye and three respondents had a weak left eye. The average visual acuity for both eyes was 1.29 with a SD of 0.38. A handedness test was also conducted with the Annett Handedness Inventory (1970), which revealed that only one participant was left handed and the remaining 16 participants were right handed. In total 11 participants were male and six participants were female. The age of the male participants ranged from 19 to 28 with a mean age of 22.5 and a SD of 2.3. Female participants ranged from 20 to 26 with a mean age of 22.8 and SD of 2.6. The majority of the participants originated from The Netherlands (N = 11) and the remaining participants originated from Germany (N = 6). All the participants in this study were students from the University of Twente and therefore have a high knowledge level of the English language.

2.2 Stimulus selection and task

The 72 website homepages used in this study were provided by Tuch et al. (2012). The stimuli were based on company website homepages as Roth et al. (2010) showed that users have developed a consistent mental model of company websites, this is important for PT measurements. The ratings between participants on VC and PT were consistent with interclass correlations of .82 for VC and .78 for PT. VC and PT were independently manipulated in the set of website homepages. Only the main effect was significant and there was no interaction between VC and PT. As these websites are selected and rated on VC and PT by external researchers, the assumption has been made that their stimulus selection process and tests are correct. More information about how the websites are selected by Tuch et al. (2012) can be found in their paper and in appendix 5.2.

Tuch et al. (2012), Lindgaard et al. (2006, 2011), Tractinsky et al. (2006) and to my knowledge all other researchers in the field of first impression website attractiveness all exposed a single website homepage at a time to the respondent which would be rated on attractiveness. In this study every stimulus consisted of two website homepages and the respondent had to decide whether they found the left of right website homepage the most attractive (e.g. figure 2 on the next page). This dual exposure gives a more direct indication of preferences in comparison with the rating scales used in other studies. Another advantage of the dual exposure is that more useful EEG data could be gathered as it allows for event-related lateralization (ERL) analysis.

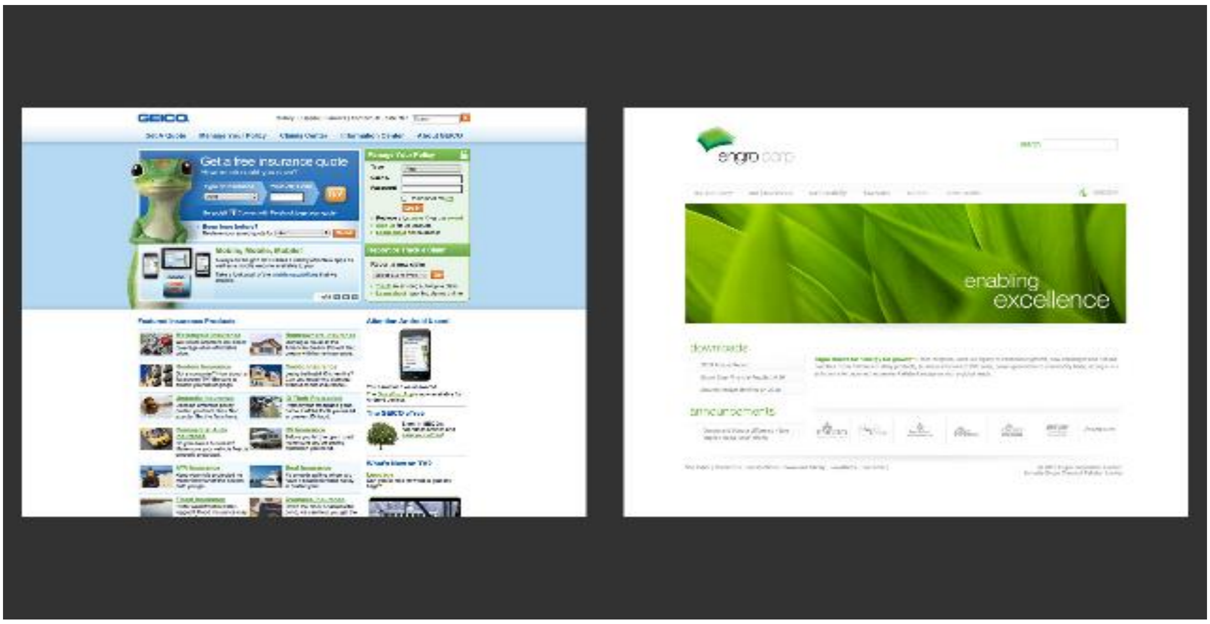


Figure 2 An example stimulus used in this study which consist of two website homepages.

The two websites that create one stimulus were always different in the degree of VC and/or PT. In total this would result in 16 (2x2x2x2) different categories. However, after removing identical categories the set was decreased to 6 categories as depicted in table 1 on the next page. For instance, the category VC high, PT low (left website) and VC and PT low (right website) was removed as it is identical to category 1 in table 1. The only difference is that the left and right website are mirrored. Each category contains 12 different website homepages. Random combinations of the website homepages were created to construct 18 stimuli for each category. In total this resulted in 72 different website homepages that constructed 108 different stimuli.

Every stimulus was shown for a total of four times due to two important factors. The first factor was the exposure duration of the stimuli. The stimuli were shown for 500ms or for 50ms. The second factor took the effect of website position into consideration. This means that every stimulus was also displayed as a mirrored version (location left and right website homepage switched). This resulted in 432 stimuli in total (108x2x2) divided between 24 categories (6x2x2). The 24 different categories are shown in appendix 5.3. All stimuli were shown in a completely random order for every participant.

	LEFT WEBSITE		**RIGHT WEBSITE**		**CATEGORY NAME**
CATEGORY	**VC**	**PT**	**VC**	**PT**	
1	Low	Low	High	Low	VLPL-VHPL
2	Low	Low	Low	High	VLPL-VLPH
3	Low	Low	High	High	VLPL-VHPH
4	High	Low	Low	High	VHPL-VLPH
5	High	Low	High	High	VHPL-VHPH
6	Low	High	High	High	VLPH-VHPH

Table 1 The different categories which all have a different combination of high/low VC and PT. The abbreviation of the categories is also given. The characters on the left side of the dash sign in the abbreviation display the stimulus properties of the left website. The right side displays the stimulus properties of the right website.

The trial of showing stimuli and participants reporting their preference is shown in figure 3. Every trial started with a white fixation point which turned red after 2000ms, the duration of the red fixation point was 500ms. Thereafter the two website homepages were exposed to the participant for either 500ms or 50ms. Every website exposure was followed by a mask to reduce the time that the website homepages were perceived. Two identical noise masks were used to mask the left and right website. This backward mask was shown after 50ms or 500ms relative to the stimulus onset, depending on the stimulus exposure duration. Finally the participants were forced to report whether they found the left or right website the most attractive. They could indicate their preference with a left or right CTRL button press. There was no time restriction on making a response and the participants were not required to respond as fast as possible.

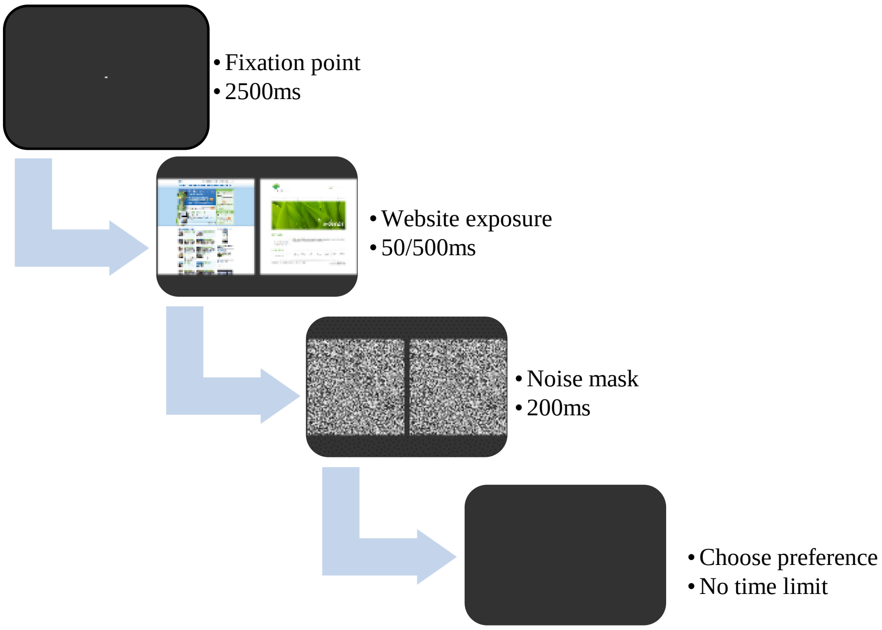


Figure 3 A representation of a single stimulus exposure.

2.3 Procedure

The experiment was conducted in the EEG lab of the Faculty of Behavioural, Management and Social sciences at the University of Twente. This lab provided a silent room with artificial lighting. Before participation every participant received an email with written instructions about the research. It was explicitly mentioned that participants were not allowed to drink alcohol 24 hours in advance of the research and that hair should be washed. Upon arrival at the EEG lab participants received more detailed information and instructions about the research. Two forms were required to be completed, an informed consent form and a participant registration form. After completion of the forms a colour-blindness and acuity test was conducted. Thereafter, the participant were seated in an office chair in front of the computer. The distance from the computer screen and the participants eyes was 85 cm and the distance from the table and the participants eyes (height) was 45 cm. Finally the cap and electrodes would be prepared.

The task could begin after the resistance of the electrodes was below 10 kiloohm and no excessive noise was apparent. Lights were turned off and curtains were closed in order to create constant exposure conditions for all participants and to reduce noise in the PCN. The experiment file displayed the instructions of the task. These instructions were read aloud by the researcher. When the instructions were clear to the participant he or she could begin the experiment. Ten practice trials were shown first in order to make sure that the participant understood their task. Thereafter, the stimuli were shown in a completely random order in 432 trials and the participant indicated whether they found the left or right website the most attractive. The average duration of the experiment ranged from 27.2 to 44.2 minutes with a mean of 32.5 minutes and a SD of 4.7 minutes. When removing outliers (participants that required tweaking of the electrodes during the experiment) the mean was 30.8 minutes with a SD of 1.95 minutes

2.4 Materials, apparatus and EEG recordings

Participants were seated in an office chair in front of a computer. The AOC G2460P LED computer screen had a diagonal diameter of 24 inch and a refresh rate of 144hz. Presentation 18.1 build 02.01.15 was installed and used to present the stimuli. A QWERTY keyboard was used on which three buttons were active: left CTRL, right CTRL and the space bar. The EEG data was recorded according to the extended 10-20 system. 32 Ag/AgCl electrodes were used and placed in an elastic cap from the brand EASYCAP (EASYCAP GmbH). The location of the 32 electrodes is as follows: AFz, AF3, AF4, AF7, AF8, Fz, F3, F4, F7, F8, FC1, FC2, FC5, FC6, Cz, C3, C4, T7, T8, CP1, CP2, CP5, CP6, Pz, P3, P4, P7, P8, POz, PO3, PO4, PO7, PO8, Oz, O1, O2. The position of the electrodes are also shown in figure 4. The ground electrode was placed on the forehead of the participant. To correct for eye movements the horizontal and vertical electro-oculogram (hEOG and vEOG) were measured. The hEOG was placed besides the left and right eye (besides outer canthi) and the vEOG was placed above and below the left eye. As mentioned, the resistance of the electrodes were kept below 10 kiloohm before the experiment would start. The resistance was decreased with SuperVisc HighViscosity Electrolyte-Gel (EASYCAP GmbH) and other standard procedures to improve the conductivity. The EEG electrodes, ground electrode, hEOG and vEOG were amplified with the actiCHamp (Brain Products GmbH). This data together with task related events (e.g. stimulus, response, mask) were sent to another computer where BrainVision Recorder version 1.21.0004 was installed (Brain Products GmbH).

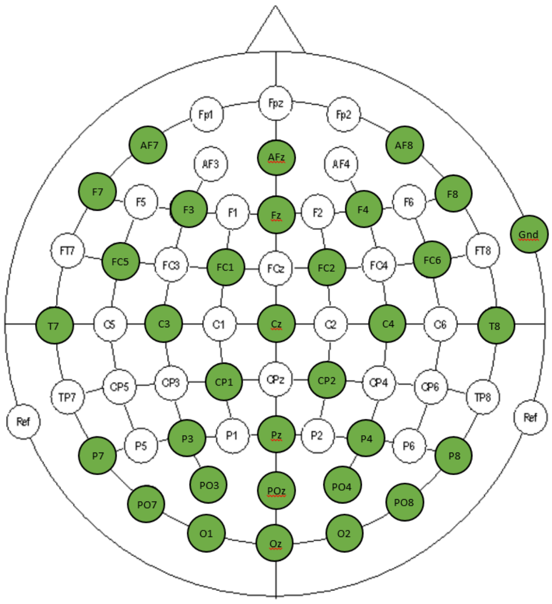


Figure 4 The green coloured circles represent the positions of the used electrodes.

2.5 Behavioural data analysis

Behavioural data was collected with Presentation 18.1 build 02.01.15. The responses and response times, amongst others, were measured for each of the 432 trials for every participant. This data including stimulus properties (e.g. category number and exposure duration) where then exported to IBM SPSS Statistics 23 where statistical analysis was conducted.

In total two two-way repeated measures ANOVA analyses (RMAAs) were conducted. This analysis included the variables category and exposure duration. In order to achieve relevant results relating to the effect of high VC and high PT it was required to compare two categories in which only the degree of VC or PT was different. Therefore it was necessary to keep the stimulus properties of the left or right website constant. For instance, a significant difference between the category VLPL-VHPH and

VHPL-VLPL cannot solely be related to either VC or PT. However, when one website category is held constant, e.g. VHPL-VLPL and VHPH-VLPL, it is possible to relate the significant difference to the low PT. As this is the only variable between these two categories.

The effect of high VC and high PT was analysed four times as there are four types of websites in this research and each type of website is used as a constant once (VHPL, VLPH, VLPL and VHPH). For example, the categories VLPL-VHPL, VLPH-VHPL and VHPH-VHPL were compared at both exposure durations. Note that VHPL is constant and the other website type varies. Each two-way RMAA that is conducted covers two constant categories. The two two-way RMAAs that were conducted are displayed in table 2 and the four constant website types are displayed behind the brackets. Category six in the table is not relevant for the effect of high VC and PT but it is relevant for the overall effect of category and time and was therefore included in this analysis.

After this analysis another two-way RMAA was conducted. This analysis also included the variables category and exposure duration. Instead of six different categories, this analysis only included two different categories. The first category was high VC and the second category was high PT. The categories were calculated by averaging the effect of VC and PT over all constant categories. As mentioned before, the combined categories VHPL-VLPL and VHPH-VLPL were compared in order to analyse the effect of PT, as this is the only variable. This is done for VC and PT, for every constant

TWO-WAY REPEATED MEASURES ANOVA ANALYSIS

CATEGORY	**Two –way RMAA 1** (Constant VHPL and VLPH)		**Two-way RMAA 2** (Constant VLPL and VHPH)	
1.	VLPL-VHPL		VHPL-VLPL	
2.	VHPH-VHPL	VHPL	VLPH-VLPL	VLPL
3.	VLPH-VHPL		VHPH-VLPL	
4.	VLPH-VLPL	VLPL	VHPH-VHPL	VHPH
5.	VLPH-VHPH		VHPH-VLPH	
6.	VHPH-VLPL		VLPH-VHPL	

Table 2 Two-way RMAAs constant categories. Note that the constant website types are displayed behind the brackets.

category and this was thereafter averaged. After this two-way RMAA a t-test against zero has been conducted on the same data in order to test whether the effect of VC and PT were significant and positive or negative.

2.6 PCN analysis

The PCN has been extracted from the EEG with Brainvision Analyzer 2.1 software (Brain Products GmbH) and was analysed with SPSS. The raw EEG data of every participant was first split into segments that each include the data from 500ms before and 3000ms after the stimulus exposure. As 432 stimuli are included in this study the same amount of segments were created. Segmentation was followed by the establishment of baselines from -100ms to 0ms relative to the stimulus exposure. Thereafter two different artefact rejections were performed. The first artefact rejection allowed a maximal voltage of 100 µV/ms, minimal and maximal amplitude of +/- 250 µV and finally the lowest activity in intervals of 0.1 µV. Bad segments were rejected based on individual channel mode. The first artefact rejection was followed by an ocular correction with the Gratton & Coles method. Both the horizontal EOG and vertical EOG were applied in the ocular correction. Next, the second artefact rejection was performed. The minimal and maximal allowed amplitude was lowered until +/- 80 µV, the other criteria remained unchanged. The remaining data was then segmented into 24 groups which represent the 24 categories mentioned before (see appendix 5.3). The average was calculated for every channel in each of the 24 segments. This calculation was followed by computing the event-related lateralization (ERL) (Van der Lubbe & Utzerath, 2013; Wascher & Wauschkuhn, 1996). In this study the ERL was calculated between the mirrored and unmirrored version of the categories. For instance, the ERL was calculated between category 1 and category 3 (see appendix 5.3). The above steps resulted in ERL measures for every individual participant which are used in statistical analysis.

Mean activity regarding the ERL data of every participant with time intervals of 20ms were exported to SPSS where statistical analysis was conducted. In total 25 timewindows were exported which included the PO7/PO8 lateralization data from 0ms towards 500ms relative to stimulus onset. The analysis of this PCN is similar to the behavioural data, as the data is analysed twice to cover all constant categories (table 2). First two three-way RMAAs were conducted with timewindow, category and exposure duration as variables. Thereafter each three-way RMAA was split into 25 two-way RMAAs which included the variables category and exposure duration. Each of these 25 two-way RMAAs covered one timewindow of 20ms long and therefore it was possible to determine when certain effects occur. When analysing a total of 25 timewindows chance will play a large role and type 1 error may come into play. The significance criterion of a single timewindow is normally set at 0.05. However, since 25 timewindows were tested, the cumulative probability of having at least one type 1 error is 72.2%. Therefore, a significant effect was only considered a significant effect when it was apparent in at least two consecutive timewindows. This reduced the cumulative probability of at least one type 1 error to only 6%. To further decrease the chance of type 1 error, an adjusted alpha level was used as described by Van der Lubbe, Bundt and Abrahamse (2014). It was computed with the following formula: $P_{crit} = \sqrt{(0.05/((\text{timewindows}-1) \times \text{electrodes}))}$. In this study, with 25 timewindows and one relevant electrode, the adjusted alpha level was < 0.0456. To make the alpha level easier to interpret an alpha level of 0.045 was used instead. This alpha level reduces the cumulative probability of at least one type 1 error to 4.9%.

After this analysis, another three-way RMAA was conducted to analyse the effect of VC and PT. This analysis included the variables timewindow, category and exposure duration. Instead of six different categories, this analysis only included two different categories. The first category is high VC and the second category is high PT. The categories were calculated by averaging the effect of VC and PT over all constant categories for every timewindow, as explained under the behavioural data analysis. Thereafter, a t-test against zero with an alpha of 0.0125 was conducted in order to test whether the effect of VC and PT were significant and positive or negative.

3. Results

3.1 Behavioural data results

In the experiment participants had to decide whether they found the left or right website the most attractive. Figure 5 below displays the mean preference of the 17 respondents for each of the six website categories. It is important to note that the mean preference indicates the percentage of respondents that preferred the first website type listed in the category name.

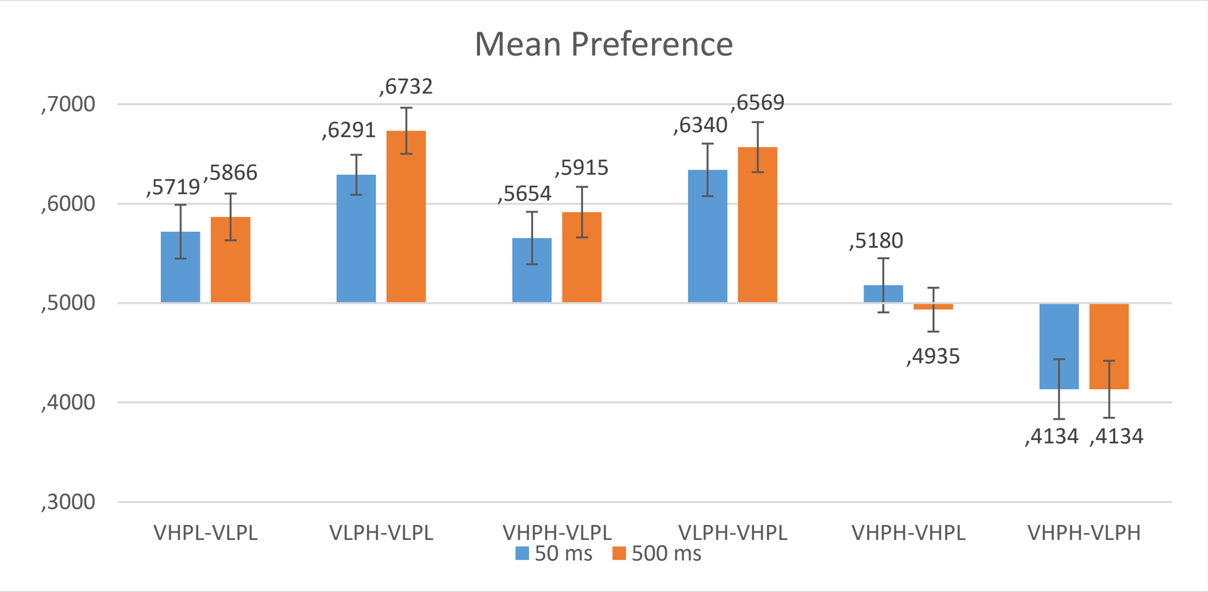


Figure 5 Mean preference graph. The figures and bars display the proportion of responses that preferred the left website type in the category name for an exposure duration of 500ms and 50ms. The level of chance is 0.5.

The mean preference data revealed multiple things regarding high VC, high PT and exposure duration. First, looking at the 500ms data, it becomes clear that the different types of websites are preferred differently. The most preferred website type is VLPH. The average preference for the VLPH category was approximately 67% when compared to VLPL websites, 66% when compared to VHPL websites and 59% when compared to a VHPH website in the 500ms condition. The categories VHPL and especially VLPL are not preferred by the participants. The preference for VLPH and the dispreference for VHPL suggests that high VC has a negative effect and that high PT has a positive effect on first impression visual appeal judgements of websites. The preference of VLPH over VHPH (59%) also suggests this negative relation between high VC and first impression visual appeal judgements of websites. On the other hand, the preference of VLPL over VHPL (41%) suggests a positive effect of high VC. The preference for VLPH over VLPL (67%) suggests a positive relation between high PT and first impression visual appeal judgements of websites. Finally, the category VHPH-VHPL does not indicate a clear effect of high PT as the results are close to level of chance.

The mean preference descriptives in figure 5 show besides data regarding the effect of the different categories on first impression visual appeal judgements also data regarding the effect of exposure duration. The mean preferences in the 500ms condition were very similar to the mean preferences in the 50ms condition. In fact, the mean preference of the different categories in the 50ms condition differed on average only 0.022 (2.2%) from the 500ms condition. The minimum difference was lower than 0.000 (0.00%) and the maximum difference was 0.0441 (4.41%). Moreover, these mean preference differences in the 50ms condition were in almost all cases closer to 50% in comparison with the 500ms condition. This seems logical as the decrease in exposure duration will increase the uncertainty of the respondents and therefore chance will play a larger role in their decisions.

Looking at the mean preference data there seems to be an effect of category, a small effect of exposure duration and an interaction effect between exposure duration and mean preference. The question is whether these effects are significant. In total two two-way RMAAs were conducted with four different constant categories as there are four different types of websites (VHPL, VLPH, VLPL and VHPH). The first two-way RMAA had the categories VLPH and VHPL as a constant and was the main analysis as they are expected to be the most and least attractive category. Comparing a category to a category that has extreme scores will likely reveal stronger results than comparing to an average category. In all analysis normality assumptions were met and therefore no correction was used on the tests of within-subjects effects, unless mentioned otherwise. Interaction results show that there was no significant two-way interaction between category and exposure duration, $F(5.80) = 1.165$, $P = 0.334$, $\eta_p^2 = 0.068$

Now it is known that there is no interaction effect we can look into the main effects. The first main effect that will be discussed is the exposure duration. Results showed that there is no significant main effect of exposure duration on first impression visual appeal judgements of websites, $F(1,16) = 0.595$, P = 0.452, η_p^2 = 0.036. This suggests that first impression visual appeal judgements of websites made in 50ms are not different from 500ms.

The second main effect that was analysed is the effect of category. The main effect of category does not meet the assumption of sphericity ($\chi2(2) = 32.071$, p = .004) and therefore the Greenhouse-Geisser correction was used. Results show that there was a significant main effect of category on first impression visual appeal judgements, $F(5,80) = 15.639$, $p < 0.0005$***, η_p^2 = 0.494. This significant main effect of category does indicate that the different categories are preferred differently, but does not show how VC and PT influence these preferences. Pairwise comparisons were conducted with an adjusted alpha level of 0.0125 as we were interested in four pairwise comparisons (0.05 / 4 = 0.0125). The pairwise comparisons reveal that the combined category VHPH-VHPL has a significantly lower mean preference when compared to VLPH-VHPL (M[2] = -0.140, SE = 0.027, 98.75% CI [-0.214, -0.065], $p < 0.0005$). The only difference between these combined categories was the level of VC and the significant results can therefore be attributed to the level of VC. In other words, high VC had a negative effect on first impression visual appeal judgements of websites. Besides an effect of VC we could also see an effect of PT. The combined category VLPH-VHPL had also a significantly higher mean preference when compared to VLPL-VHPL (M = 0.225, SE = 0.034, 98.75% CI = [0.130, 0.320], $p < 0.0005$). This provided evidence for PT having a positive effect on first impression visual appeal judgements of websites. So far only one two-way RMAA and only the effect of VC and PT when VHPL is kept constant are discussed. In table 3 on page 26 the results are displayed when the other website categories were used as a constant. The results are similar but in some cases these results are not significant. First no significant interaction effect could be found between category and exposure duration. Second, the main effect of category was significant in both the RMAAs and the main effect of exposure duration was not significant in any of the analysis. Moreover, the negative effect of high VC was highly significant in two of the four analysis and the positive effect of high PT was also highly significant in two of the four analysis.

[2] M indicates mean difference from 0.

Table 3 on page 26 displays the effect of VC and PT for every constant category. To correctly represent the variables VC and PT, all pairwise comparisons were averaged for every participant. A two-way RMAA with the variables category (VC high and PT high) and exposure duration was conducted. This analysis revealed that there was a highly significant main effect of category, $F(1,16) = 43.582$, $P < 0.0005$, $\eta_p^2 = 0.731$. This effect of category did not interact with exposure duration, $F(1,16) = 0.731$, $P < 0.405$, $\eta_p^2 = 0.044$. Pairwise comparison revealed that high VC was preferred significantly lower than high PT, (M = -0.145, SE = 0.022, 95% CI [-0.192, -0.099], p < 0.0005). To determine whether high VC and high PT have a positive or negative effect on visual appeal judgements of websites, a t-test was conducted which tested against zero. This test revealed that high VC has a negative effect on first impression visual appeal judgements of websites, (t(16) = -2118, P = 0.050, M = -0.033, SE = 0.016, 95% CI = [-0.067, 0.000]). This result is barely significant even though, as mentioned before, two of the four constant categories show highly significant results for VC. This is due to the fact that the constant category VHPH revealed a positive effect of VC that is close to being significant (see table 4), which has a major effect on the average effect of VC. High PT, on the other hand, had a positive effect on first impressions visual appeal judgements of websites, (t(14) = 7.519, P < 0.0005, M = 0.112, SE = 0.015, 95% CI = [0.080, 0.143]).

	CONSTANT CATEGORY	INTERACTION CATEGORY X EXPOSURE DURATION	MAIN EFFECT CATEGORY	MAIN EFFECT EXPOSURE DURATION	EFFECT HIGH VC	EFFECT HIGH PT
TWO-WAY RMAA 1	VHPL	F(5.80) = 1.165, P = 0.334, η_p^2 = 0.068	F(5,80) = 15.639, p < 0.0005***, η_p^2 = 0.494 Sphericity assumption not met: χ2(2) = 32.071, p = .004	F(1,16) = 0.595, p = 0.452, η_p^2 = 0.036	M = -0.140 SE = 0.027, 98.75% CI = [-0.214, -0.065], p < 0.0005***	M = 0.225, SE = 0.034, 98.75% CI = [0.130, 0.320], p < 0.0005***
	VLPH	The same as VHPL	The same as VHPL	The same as VHPL	M = 0.006, SE = 0.021, 98.75% CI = [-0.053, 0.065], p = 0.789	M = 0.059, SE = 0.026, 98.75% CI = [-0.015, 0.132], p = 0.039
TWO-WAY RMAA 2	VLPL	F(5.80) = 0.878, P = 0.500, η_p^2 = 0.052	F(5,80) = 18.510, p < 0.0005***, η_p^2 = 0.536 Sphericity assumption not met: χ2(2) = 36.154, p = .001	F(1,16) = 1.928, p = 0.184, η_p^2 = 0.108	M = -0.073, SE = 0.020, 98.75% CI = [-0.129, -0.017], p = 0.002***	M = -0.001, SE = 0.028, 98.75% CI = [-0.081, 0.079], p = 0.977
	VHPH	The same as VHPH	The same as VHPH	The same as VHPH	M = 0.073, SE = 0.029, 98.75% CI = [-0.008, 0.154], p = 0.022	M = 0.165, SE = 0.026, 98.75% CI = [0.093, 0.238], p < 0.0005***

Table 3 Behavioural data: two two-way RMAA results for different constant categories including the effect of high VC and high PT. The first two-way RMAA included the constant categories VHPL and VLPH. The second two-way RMAA included the constant categories VLPL and VHPH.

3.2 PCN results

The average PCN of the respondents for the different website types is shown in figures 6 – 11 and reveals interesting information. It is important to note that a negative PO7 indicates a preference for a website category and not a positive PO7. The figures reveal that the PO7/PO8 lateralization started to be visible during or after the 80 – 100ms timewindow. Moreover, the 50ms condition seemed to show similar PO7/PO8 lateralization to the 500ms condition from approximately 80 – 140ms. This is not the case for the combined category VLPH-VHPL, which showed a weak PO7/PO8 lateralization in general in this timewindow. Statistical analyses on the hEOG revealed that there were no significant eye movements when displaying the different website categories. The first three-way RMAA revealed that there were a couple of timewindows where the hEOG was significant ($F(25,350) = 2.449$, $P < 0.0005$, $\eta_p^2 = 0.149$). In the methodology it was mentioned that a significant effect of PO7 or the hEOG only was significant when it occurred in two consecutive timewindows with an alpha level of 0.045, in order to reduce type 1 error. Testing the 25 timewindows with time intervals of 20ms revealed that the eye movements that were significant did not appear in two consecutive timewindows and can therefore be regarded as not significant. The second two-way RMAA revealed no significant eye movements over time ($F(25, 450) = 1.466$, $P = 0.072$, $\eta_p^2 = 0.095$).

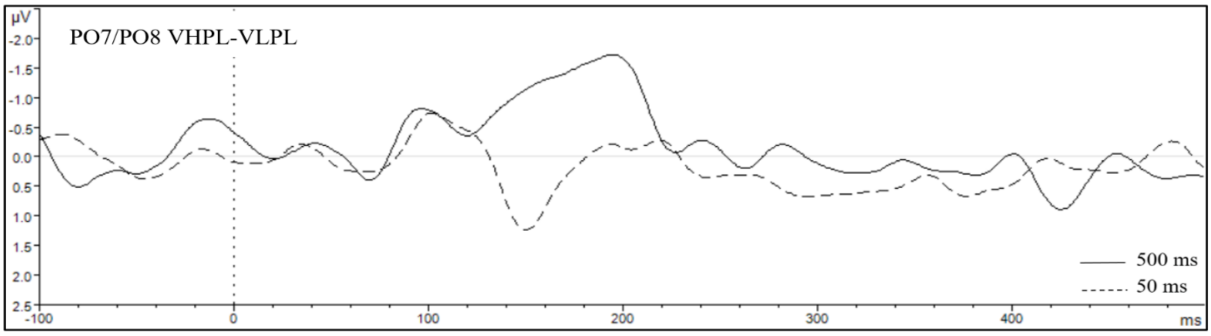


Figure 6 PO7/PO8 lateralization for combined category VHPL-VLPL. Note that the straight line indicates the 500ms condition and the dotted line the 50ms condition.

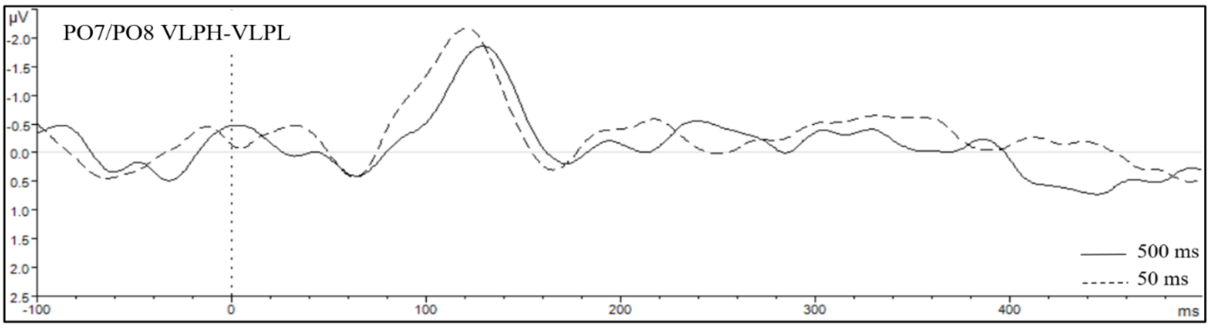


Figure 7 PO7/PO8 lateralization for combined category VLPH-VLPL. Note that the straight line indicates the 500ms condition and the dotted line the 50ms condition.

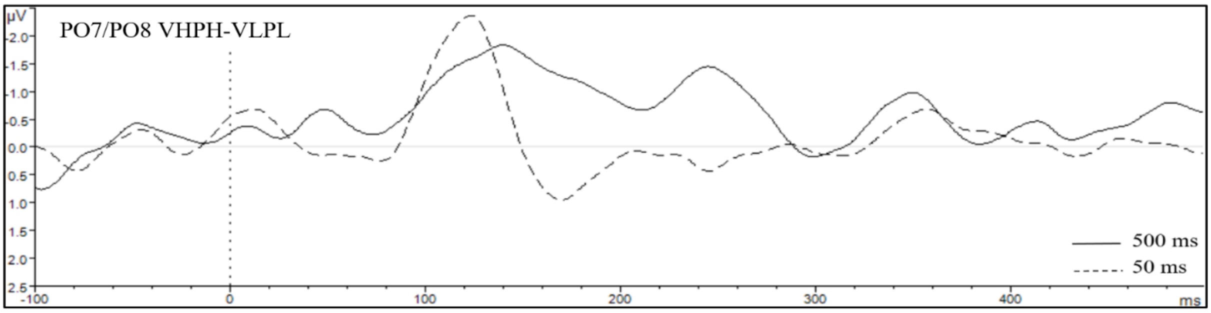


Figure 8 PO7/PO8 lateralization for combined category VHPH-VLPL. Note that the straight line indicates the 500ms condition and the dotted line the 50ms condition.

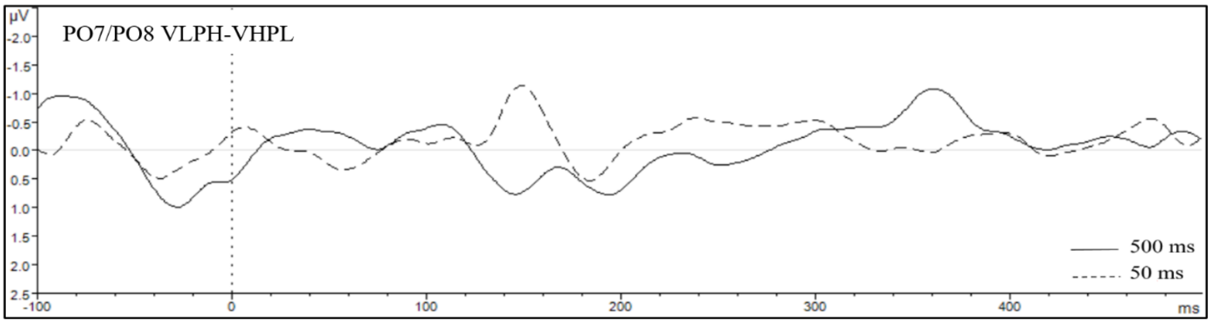


Figure 9 PO7/PO8 lateralization for combined category VLPH-VHPL. Note that the straight line indicates the 500ms condition and the dotted line the 50ms condition.

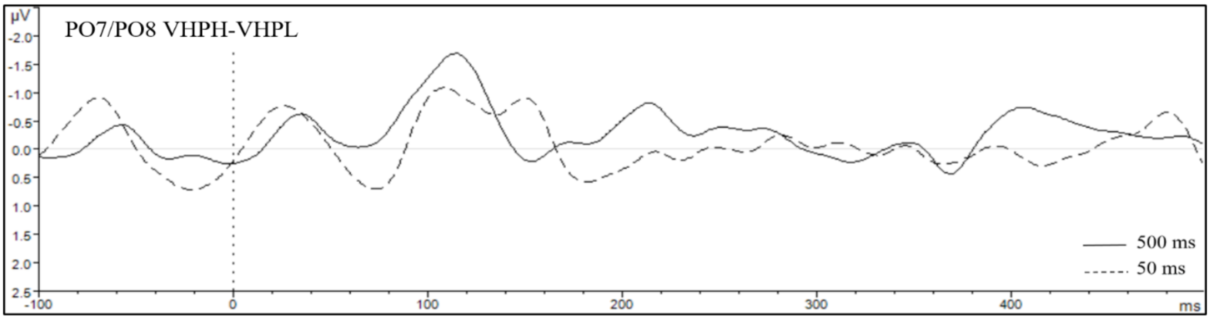


Figure 10 PO7/PO8 lateralization for combined category VHPH-VHPL. Note that the straight line indicates the 500ms condition and the dotted line the 50ms condition.

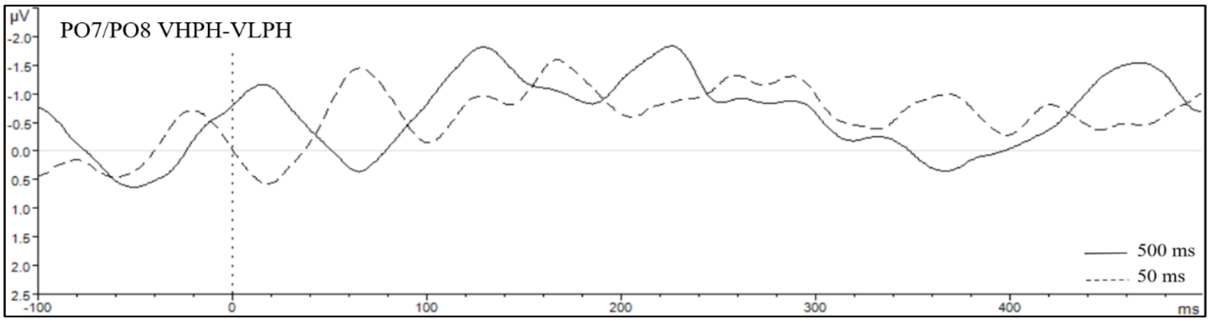


Figure 11 PO7/PO8 lateralization for combined category VHPH-VLPH. Note that the straight line indicates the 500ms condition and the dotted line the 50ms condition.

Similar to the analysis of the behavioural data, two three-way RMAAs were conducted for the PCN. These three-way RMAAs show the main effect of timewindow, category, exposure duration and their interactions. As the results of both three-way RMAAs were very similar, only the first three-way RMAA results with constant category VHPL and VLPH are discussed. All results are displayed in table 5 on the next page. First, and most important, we could see a highly significant effect of timewindow, F(25.350) = 2.938, P < 0.0005, η_p^2 = 0.173. This indicates that the measured PO7 electrode changed over time. The effect of category was also highly significant F(5.70) = 4.741, P = 0.001, η_p^2 = 0.253. This indicates that the different categories resulted in different PO7 values and therefore indicate different preferences. How these different preferences relate to VC and PT will be discussed later. The final main effect, which is exposure duration, was not significant, F(1.14) < 0.0005, P = 0.996, η_p^2 < 0.0005. The two-way interaction of timewindow and category was significant, F(125.1750) = 2.150, P < 0.0005, η_p^2 = 0.133. So as one could expect, the effect of category on PO7 was stronger in certain timewindows. The two-way interaction of timewindow and exposure duration was not significant, F(25.350 = 1.396, P = 0.101, η_p^2 = 0.091). The final two-way interaction of category and exposure duration was also not significant, F(5.70) = 0.849, P = 0.520, η_p^2 = 0.057. Finally the three-way interaction of timewindow, category and exposure duration was highly significant, F(125.1750) = 1.444, P = 0.001, η_p^2 = 0.093. This shows that the effect of exposure duration was stronger for some categories and was also dependant on the timewindows.

	TIMEWINDOW (T)	CATEGORY (C)	EXPOSURE DURATION (D)	T*C	T*D	C*D	T*C*D
TWO-WAY RMAA 1 (VHPL &VLPH)	F(25.350) = 2.938, P < 0.0005***, η_p^2 = 0.173	F(5.70) = 4.741, P = 0.001***, η_p^2 = 0.253	F(1.14) < 0.0005, P = 0.996, η_p^2 < 0.0005	F(125.1750) = 2.150, P < 0.0005***, η_p^2 = 0.133	F(25.350 = 1.396, P = 0.101, η_p^2 = 0.091	F(5.70) = 0.849, P = 0.520, η_p^2 = 0.057	F(125.1750) = 1.444, P = 0.001***, η_p^2 = 0.093
TWO-WAY RMAA 2 (VLPL & VHPH)	F(25.350) = 5.492, P < 0.0005***, η_p^2 = 0.282	F(5.70) = 2.124, P = 0.073, η_p^2 = 0.132	F(1.14) = 0.489, P = 0.496, η_p^2 = 0.034	F(125.1750) = 1.600, P < 0.0005***, η_p^2 = 0.103	F(25.350 = 0.871, P = 0.647, η_p^2 = 0.059	F(5.70) = 0.704 P = 0.622, η_p^2 = 0.048	F(125.1750) = 1.542, P < 0.0005***, η_p^2 = 0.099

Table 5 Two three-way RMAA results for different constant categories. The first two-way RMAA included the constant categories VHPL and VLPH. The second two-way RMAA included the constant categories VLPL and VHPH.

In total, 25 simple two-way RMAAs were conducted to analyse in which timewindows the main effect of category, the main effect of exposure duration and the two-way interaction of category and exposure duration occurred. Each two-way RMAA covered one timewindow with a range of 20ms and all the results are displayed in appendix 5.4. The most important results will be discussed in the following sections.

3.2.1 Interaction effect category x exposure duration

As shown in figure 12 on the next page and appendix 7.4 there was a significant interaction effect of category and exposure duration in both subsets of 25 two-way RMAAs that were conducted. The first subset of 25 RMAAs with constant categories VHPL and VLPH revealed that the interaction starts from the 140 – 160ms timewindow (F(5.70) = 7.233, P < 0.0005, η_p^2 = 0.341) and remains significant until the 200 – 220ms timewindow (F(5.70) = 2.719, P = 0.027, η_p^2 = 0.163). Similarly, the second subset of 25 RMAAs with constant categories VLPL and VHPH revealed that the interaction was significant from the 140 – 160ms timewindow (F(5.70) = 7.114, P < 0.0005, η_p^2 = 0.337) until the 180 – 200ms timewindow (F(5.70) = 2.935, P < 0.018, η_p^2 = 0.137). These significant effects were however not relevant as it occurred 40 – 60ms after the main effect of category. It is logical that there is an interaction effect between category and exposure duration after 40 – 60ms. This due to the fact that the short exposure stimuli, which were displayed for 50ms, have already disappeared. Therefore the relevant data should appear in the first three initial consecutive timewindows where category is significant. These timewindows were 80 – 140ms for the first subset of 25 RMAAs and 100 – 140ms for the second subset of 25 RMAAs. There were however no significant interaction effects in any of these timewindows. The data of the 100 – 120ms timewindow for example provided the following significance data for the first subset of RMAAs (F(5.70) = 0.773, P < 0.486, η_p^2 = 0.052) and second subset of RMAAs (F(5.70) = 0.841, P < 0.465, η_p^2 = 0.057).

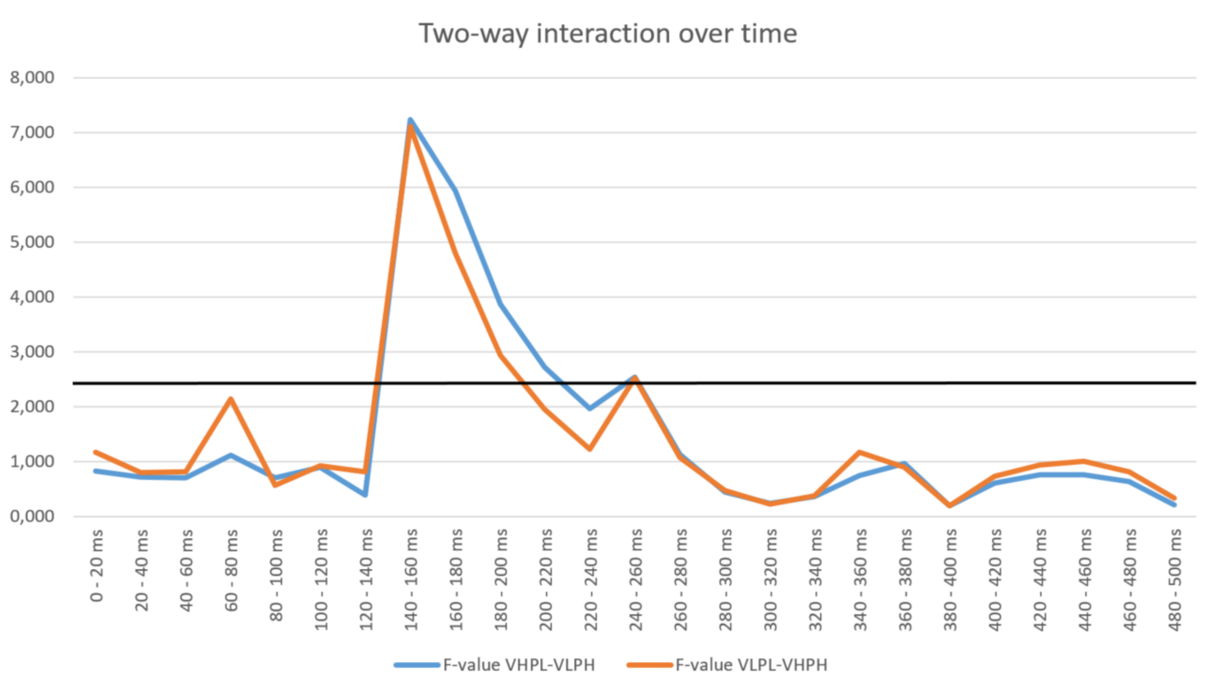


Figure 12 F-values of the interaction effect of category and exposure duration over time. All F-values above 2.408 are significant on the 0.045 level and this is represented by the black line in the figure. Note that when the Greenhouse-Geisser correction was used, the F-value in the table is altered to correctly represent the significance level of that timewindow.

3.2.2 Main effect of exposure duration

As shown in figure 13 and appendix 7.4 there was a main effect of exposure duration in the second subset of 25 two-way RMAAs with VLPL and VHPH as a constant category. This effect of exposure duration started in the 160 – 180ms timewindow (F(1.14) = 6.098, P = 0.027, η_p^2= 0.303) and continued towards the 200 – 220ms timewindow (F(1.14) = 4.986, P = 0.042, η_p^2= 0.263). This significant effect was, similar to the interaction effect between category and exposure duration, not relevant as it occurred 60ms after the main effect of category. The relevant timewindows (80 – 140ms) did not reveal a significant main effect of exposure duration.

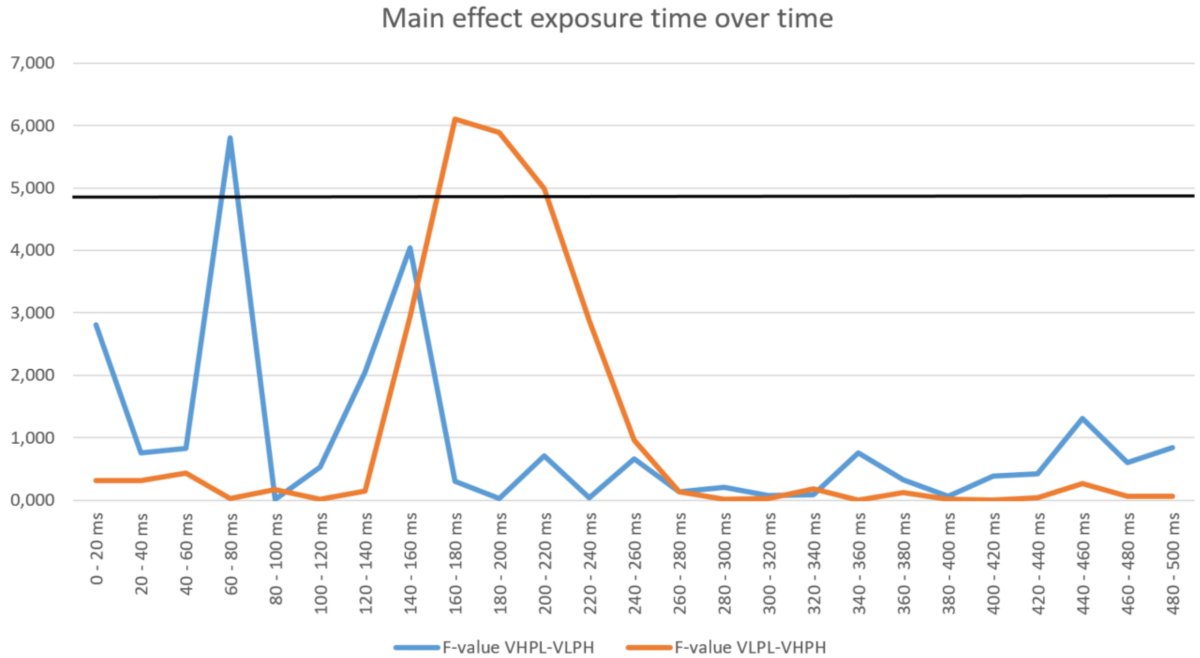


Figure 13 F-values of the main effect of exposure duration over time. All F-values above 4.845 are significant on the 0.045 level and this is represented by the black line in the figure. Note that when the Greenhouse-Geisser correction was used, the F-value in the table is altered to correctly represent the significance level of that timewindow.

3.2.3 Main effect category

As shown in figure 14 on the next page and appendix 7.4 there was a significant main effect of category in both subsets of 25 two-way RMAAs that were conducted. The figure reveals that category had a significant main effect after 80 – 100ms relative to the stimulus onset (F(5.70) = 3.061, P = 0.015, η_p^2 = 0.179) in the first subset of 25 two-way RMAAs (VLPL & VHPH constant) after a Greenhouse-Geisser correction. This main effect of category increased in significance, after a Greenhouse-Geisser correction, in the 100 – 120ms timewindow (F(5.70) = 12.451, P < 0.0005, η_p^2 = 0.471) and 120 – 140ms timewindow (F(5.70) = 17.410 , P < 0.0005, η_p^2 = 0.554). The main effect of category remained significant until the 280 – 300ms timewindow but decreased in strength. Finally, there was a significant main effect of category in the 440 – 480ms timewindow. The second subset of 25 two-way RMAAs (VHPH & VLPL constant) showed data that is similar to some extent. The significance of the main effect of category was not significant in the 80 – 100ms timewindow (F(5.70) = 0.196, P = 0.963, η_p^2= 0.014) as opposed to the first RMAA. The effect of category started however in the 100 – 120ms timewindow (F(5.70) = 3.183, P = 0.12 , η_p^2=0.185) and continued towards the 120 – 140ms timewindow, which was corrected for sphericity with the Greenhouse-Geisser correction, (F(5.70) = 5.541, P = 0.002, η_p^2= 0.284). The Timewindow 140 – 160ms was, unlike the first RMAA, not significant (F(5.70) = 1.781, P = 0.128, η_p^2= 0.113). The two following timewindows, 160 – 180ms (F(5.70) = 3.104, P = 0.014, η_p^2= 0.0.181) and 180 – 200ms (F(5.70) = 2.769, P = 0.024, η_p^2= 0.165), were significant.

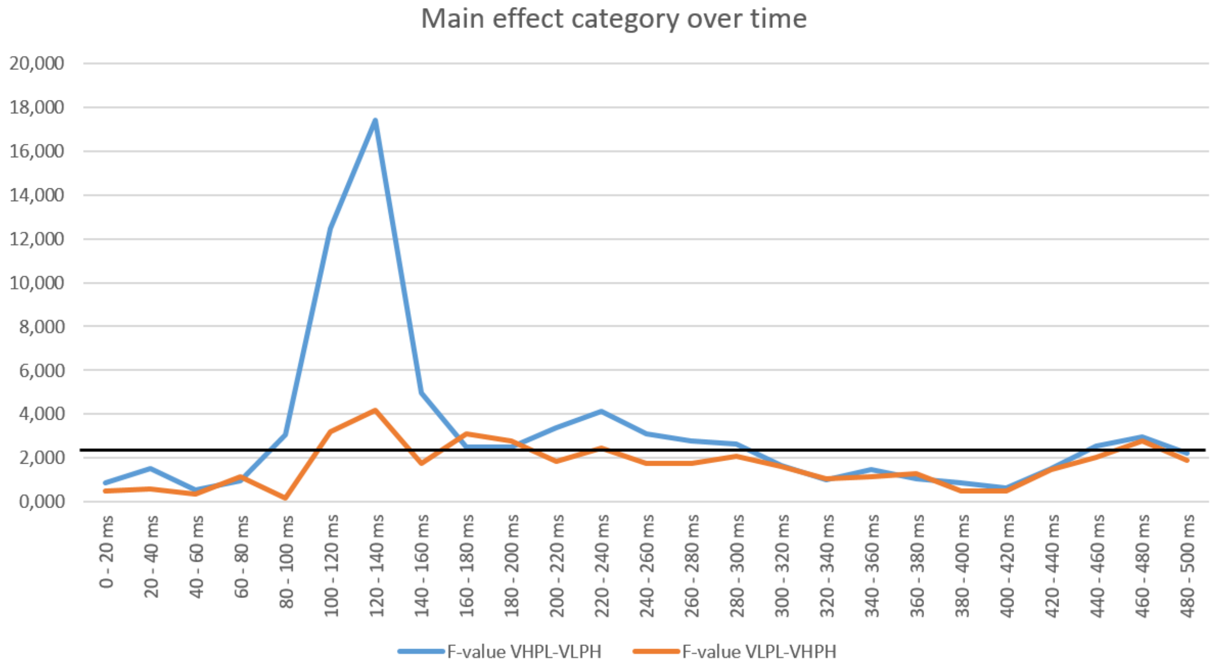


Figure 14 F-values of the main effect of category over time. All F-values above 2.408 are significant on the 0.045 level and this is represented by the black line in the figure. Note that when the Greenhouse-Geisser correction was used, the F-value in the table is altered to correctly represent the significance level of that timewindow.

There was a main effect of category on first impression visual appeal judgements of websites, but it is not yet clear how this was influenced by VC and PT. To analyse this effect of VC and PT, two separate analysis were conducted. The first analysis was a three-way RMAA with the variables timewindow, category (high VC and high PT) and exposure duration. This analysis revealed that the effects of VC and PT on PO7 were not significantly different (F(1.14) = 0.790, P = 0.389, η_p^2 < 0.053). There was also no significant two-way interaction between category and timewindow (F(24.336) = 0.656, P = 0.892, η_p^2 < 0.045) and category and exposure duration (F(1.14) = 0.089, P = 0.769, η_p^2 < 0.006). The three-way interaction of timewindow, category and exposure duration was also not significant, F(24.336) = 1.161, P = 0.276, η_p^2 < 0.077. Hereafter 25 two-way RMAAs were conducted, one for every timewindow. A significant intercept was apparent in the 100 – 120ms timewindow (F(1.14) = 23.618, P < 0.0005, η_p^2 < 0.628) and in the 120 – 140ms timewindow (F(1.14) = 25.561, P < 0.0005, η_p^2 < 0.646).

This resulted in the second analysis being conducted, a t-test in order to assess whether the PO7 electrode was significantly different from zero in every timewindow for websites with high VC and websites with high PT. It is important to note that a negative PO7 indicates a preference for a website category and not a positive PO7. As shown in the main effect of exposure duration and the interaction effect of category and exposure duration, there was no significant difference between the 500ms and the 50ms condition. The t-tests were therefore conducted on the average data of both exposure duration conditions with an adjusted alpha level of 0.0125. The results showed that there was a positive effect of high VC on PO7 that started in the 100 – 120ms timewindow ($t(14) = 3.615$, $P = 0.003$, $M = 0.635$, $SE = 0.176$, 98.75% CI = [0.132, 1.138]) and continued until the 120 – 140ms timewindow ($t(14) = 4.426$, $P = 0.001$, $M = 0.932$, $SE = 0.210$, 98.75% CI = [0.328, 1.534]). This positive effect of high VC, which indicates that low VC websites gain more visual attention, was similar to the behavioural data, which indicates that low VC websites are more attractive. The effect of high PT on the other hand was dissimilar to the behavioural data. The behavioural data indicated that high PT websites were more attractive. PCN however showed that PO7 was significantly higher than zero, which indicates that low PT websites gain more visual attention. This effect started being significant in the 100 – 120ms timewindow ($t(14) = 4.124$, $P = 0.001$, $M = 0.919$, $SE = 0.223$, 98.75% CI = [0.281, 1.566]) and continued until the 120 – 140ms timewindow ($t(14) = 3.098$, $P = 0.008$, $M = 0.914$, $SE = 0.295$, 98.75% CI = [0.069, 1.76]), which was similar to the effect of VC.

In short, websites with high VC and websites with high PT gain similar amounts of visual attention as they are not significantly different. This effect was visible in every timewindow and the effect of VC is therefore as fast as the effect of PT on PO7. There was a significant positive effect of high VC on PO7, which suggests that websites with low VC gain more visual attention. The behavioural data revealed that low VC websites were also perceived as more attractive compared to high VC websites. There was also a significant positive effect of high PT on PO7, which suggests that websites with low PT gain more visual attention. Low PT websites were however not perceived as attractive according to the behavioural data. Both the effects of VC and PT were apparent in the 100 – 140ms timewindows.

4. Discussion

This study provides the results to answer three main questions. Due to the fact that attractiveness as a concept is abstract and hard to measure, we examined two important factors that are linked to attractiveness. These factors are VC and PT and therefore the first question to be answered was how high VC and high PT influence first impression visual appeal judgements of websites. The second question was how this influence of high VC and high PT relates to the exposure duration of a website. In other words, are the effects of high VC and high PT stronger or weaker in a 500ms condition compared to a 50ms condition. This question was important for two reasons. The first reason was that this question could assess whether well-founded first impression attractiveness judgements can be made within an exposure duration of 50m. Other researchers already tried to assess whether 50ms is enough to make these judgements. However, as doubts may be raised regarding the control of afterimages in the majority of studies in this field, it was important to replicate these studies. (e.g. Lindgaard et al. 2006,. 2011). The second reason was that this question could assess when VC and PT start having an influence on attractiveness judgements of websites. The final question was whether there is a discrepancy between attractiveness and visual attention when judging websites on attractiveness that differ in the degree of VC and PT. Physical responses of the respondents revealed their preferences regarding attractiveness. The visual attention of the respondents was measured with EEG and EOG.

At the start of this study multiple predictions were made as can be seen in paragraph 1.6. To summarize, 1) VC was expected to have a negative effect on visual appeal judgements of websites even with an exposure duration of 50m. This was because stimuli with a low amount of VC are easier to process as they contain less information which needs to be processed. This may lead to a more positive judgement according to the fluency effect (Reber et al., 2004). 2) PT on the other hand was expected to have a positive effect on visual appeal judgements, even within 50ms, as prototypical stimuli are easier to process than atypical stimuli (Winkielman et al., 2006). 3) It was expected that the effect of high VC and high PT is stronger in the 500ms condition when compared to the 50ms condition. A shorter exposure duration will increase the uncertainty of the respondent regarding their judgements and chance is therefore expected to play a larger role in the 50m condition. 4) The expectations above also indicate that it was expected that individuals can make well-founded first impression visual appeal judgements of websites within 50ms. This was due to the fact that there will only be an effect of high VC and/or high PT in the 50ms condition when it is possible to make attractiveness judgements in 50ms. The reason for this expectation was that that emotion research has already proven that first impressions of persons can be made within 50ms. 5) VC is a perception and PT requires memory to some extent, therefore it was expected that the effect of PT on attractiveness judgements will be slower than VC. 6) Finally, it was expected that websites that are perceived as attractive also receive more visual attention. This means that high VC has a negative effect and that high PT has a positive effect on visual attention when perceiving websites. This was because emotion research has indicated that attractive individuals/faces receive more visual attention than unattractive or average individuals/faces (Maner et al., 2003; Seidman and Millar, 2013; Van Hoof et al., 2010). Similar results were expected for when websites are perceived instead of people.

4.1 VC and PT regarding attractiveness

The first hypothesis that will be discussed is: *'Visual complexity has a negative effect on attractiveness judgements of websites even within 50ms'.* Behavioural data suggested that there was indeed a negative significant effect of high VC on first impression website attractiveness even within 50ms. This result was expected and is in line with the findings from Tuch et al. (2011, 2012). Reber et al. (2004) has given a possible explanation for why low VC is perceived as more attractive than high or medium VC. The researchers propose that the more fluent a stimulus can be processed, the more attractive they are perceived. They call this the fluency effect. Websites with low VC contain less information and could therefore be processed more fluently, resulting in higher attractiveness. The negative effect of VC suggests that web designers, organizations and other individuals with a website should limit the visual complexity of their websites. To achieve a good first impression, websites should have a high visual appeal and not overload the users with information.

The second hypothesis that will be discussed is: '*High prototypicality has a positive effect on attractiveness judgements of websites even within 50ms'.* The behavioural data in this study has shown that there is a significant and positive effect of high PT on first impression visual appeal judgements of websites. Indicating that websites with high PT are perceived as more attractive. This result was expected and is in line with the research of Tuch et al. (2012). The fluency effect of Reber et al. (2004) could also explain why high PT websites are preferred. Winkielman, Halberstadt, Fazendeiro and Catty (2006) have found that prototypical stimuli are easier to process than atypical stimuli. According to the fluency effect this easier processing will result in higher perceived attractiveness. In practice, websites should meet the mental model that their visitors have built of (company) websites. Roth et al. (2010) revealed that web users have certain expectations related to the position of web objects (e.g. search field and navigation). The location of many objects, but not all objects, is agreed upon by the web users. Meeting this expectation will enhance the websites' PT and therefore visual appeal is also positively affected.

The third hypothesis that will be discussed is: '*The effect of high visual complexity and high prototypicality on attractiveness judgements of websites is stronger in the 500ms condition in comparison to the 50ms condition'*. The behavioural data indeed showed that the preferences for a website type in the 50ms condition were in almost all cases closer to the level of chance in comparison with the 500ms condition. Significance tests revealed however that this difference between the two conditions was not significant. The third hypothesis is therefore rejected. This result was not as expected as it was predicted that a shorter exposure duration would increase the uncertainty of the respondents regarding their judgements. Higher uncertainty was expected to increase the level of chance in the respondents' judgements and therefore decrease the strength of both VC and PT. It is important to note that more respondents or a bigger set of stimuli might reveal that there is indeed a significant effect of exposure duration on visual appeal judgements of websites. The finding that the 50ms and 500ms condition are similar was also found in other studies. For instance, Lindgaard et al. (2006, 2008) showed that the attractiveness ratings in 50ms were similar to the ratings in 500ms. The results of Tuch et al. (2012) were also similar, but they were related to VC and PT. In short, their study showed that there was no interaction between VC and exposure duration (50ms vs. 500ms vs. 1000ms) and between PT and exposure duration.

The fourth hypothesis is: '*The effect of high visual complexity on attractiveness judgements of websites is faster than the effect of high prototypicality'.* The behavioural data indicated that the effect of VC and PT are not significantly stronger or weaker in the 50ms condition compared to the 500ms condition. This indicates that 50ms is enough time to achieve the full effect of VC and PT. Therefore the fourth hypothesis is rejected based on the results of this study. This result was different than initially expected. It was expected that the effect of VC would be faster than the effect of PT, as VC is a perception and PT requires memory to some extent as individuals have to relate to the mental models that they have built of website designs. Moreover VC is included in the perceptual analysis which is the first processing stage that is involved with making attractiveness judgements according to Leder et al. (2004). PT on the other hand is included in the second processing stage. As these stages occur sequential when making attractiveness judgements, the effect of VC should appear before VC. Tuch et al. (2012) conducted a similar experiment, but applied shorter exposure durations of 17ms, 33ms and 50ms in the second experiment of their study. The results indicated that the effect of PT is weaker than the effect of VC in a short exposure duration of 17ms and gained strength equal to VC when the exposure duration was increased towards 33ms. The results in this study support the explanation given by Tuch et al. (2012) that an exposure duration of 50ms may already be enough to reach the implicit memory integration stage of the information-processing stage model of aesthetic processing (Leder et al., 2004), which includes the effect of PT. This means that 50ms is already enough to achieve the full effect of PT.

4.2 Website attractiveness judgements in a short exposure

Hypothesis one, two and three also reveal that well-founded first impression visual appeal judgements can be made in an exposure duration of only 50ms and that these judgements are similar to a 500ms exposure. The different categories of websites (VHPL, VLPH, VHPH, VLPL) in this study were preferred significantly different in the 50ms condition and did not interact with exposure duration. These findings are similar to the findings from previous studies which indicated that 50ms is enough to make a well-founded first impression of a website (e.g. Lindgaard et al., 2006, 2011; Tuch et al., 2012). As mentioned before, doubts may be raised regarding the control of afterimages in the majority of these studies. Results in this study do however reveal that even with a mask, similar results are retrieved. According to these findings we can expect that there is already a halo effect after 50ms when visiting websites. This is due to the fact that first impressions are critical for making a more thorough opinion and can influence subsequent experiences (Lindgaard et al., 2006, 2011; Lorenzo-Romero et al., 2013; Van Schaik & Ling, 2009).

4.3 Discrepancy between attractiveness and attention

The final hypothesis is: '*There is no discrepancy between attractiveness judgements of websites and visual attention for high/low VC and/or PT websites'*. Behavioural data suggested that high VC has a negative effect on first impression visual appeal judgements of websites. The analysed activity regarding the PO7 electrode was significantly higher for websites with high VC, indicating that websites with high VC are not only perceived as less attractive but also receive less visual attention. The effect of VC on PO7 was visible in the 100 – 140ms timewindow. Behavioural data regarding PT revealed that high PT has a positive effect on first impression visual appeal judgements of websites. The PCN did reveal however that websites with high PT receive less visual attention, as the measured PO7 electrode was significantly lower for websites with low PT in the 100 – 140ms timewindow. Although there was an effect of high VC and high PT on visual attention, participants were not actively looking at the different websites. This is due to the fact that the EOG revealed no significant eye-movements. The finding that high PT websites receive less visual attention but are perceived as more attractive, reveals that not only attractiveness is important for drawing visual attention. An explanation for the different results between the behavioural data and the PCN could be that atypical websites attract a lot of visual attention due to them not being standard (Schoormans & Robben, 1997; Van Hoof, Crawford and Van Vugt, 2010). However, once the attention has moved to the atypical website, people realize that this abnormal website is in fact not attractive, but only differ from their mental image that they have of company websites. There are also other variables that could influence visual attention, for instance size and colour (Schoormans & Robben, 1997). Another explanation could therefore be that the underlying variables of PT result in low PT drawing visual attention. To give an example, low PT websites could make use of certain colours, contrasts or pictures sizes that draw visual attention. The underlying variables of VC and PT were however, similar to Tuch et al. (2012), not yet fully understood and controlled for in this study. Therefore we cannot examine how PT relates to these variables.

4.4 Limitations and future research

This study has several important limitations and asks for future research. The first limitation is that the external validity may be limited as the stimuli in this study were all images of company websites. Roth et al. (2010) showed that users have developed consistent mental models of company websites, which is important for PT measurements. Other types of websites are social networking sites, search engines, online shops, online newspapers, etc. These types of websites may reveal different results and the effects of VC and PT may be more or less apparent. This is something that future researchers could investigate.

Another limitation is that this study failed to find a significant difference between attractiveness judgements of websites when exposed for 50ms and 500ms. This could be attributed to the limited number of respondents and limited number of stimuli. 17 respondents participated in this research and 432 stimuli that were exposed to each respondent. The mean preference data (figure 5) indicated that the judgements in the 50ms condition were closer to the level of chance in comparison to the 500ms condition. Significance results revealed that this difference was not significant or close to being significant. However, regarding type 2 error, it is important to mention that a higher number of respondents and/or a higher amount of stimuli might result in a significant difference between both exposure duration conditions.

Finally, there are two topics for future research. The first topic for future research should focus on the underlying variables of VC and PT. The underlying variables of VC and PT were not yet fully understood and controlled for in this study and in the study of Tuch et al. (2012). An understanding of the underlying variables of VC and PT could assist web designers. The second topic for future research should focus on the effect of a short exposure duration of a website (e.g. 50ms) on buying behaviour. Lorenzo-Romero et al. (2013) conducted a study that tried to find the effect of a short exposure (1 second) on buying intentions and quality perception, but failed to find an significant effect. More research regarding this topic may reveal an exposure duration threshold for influencing this buying behaviour.

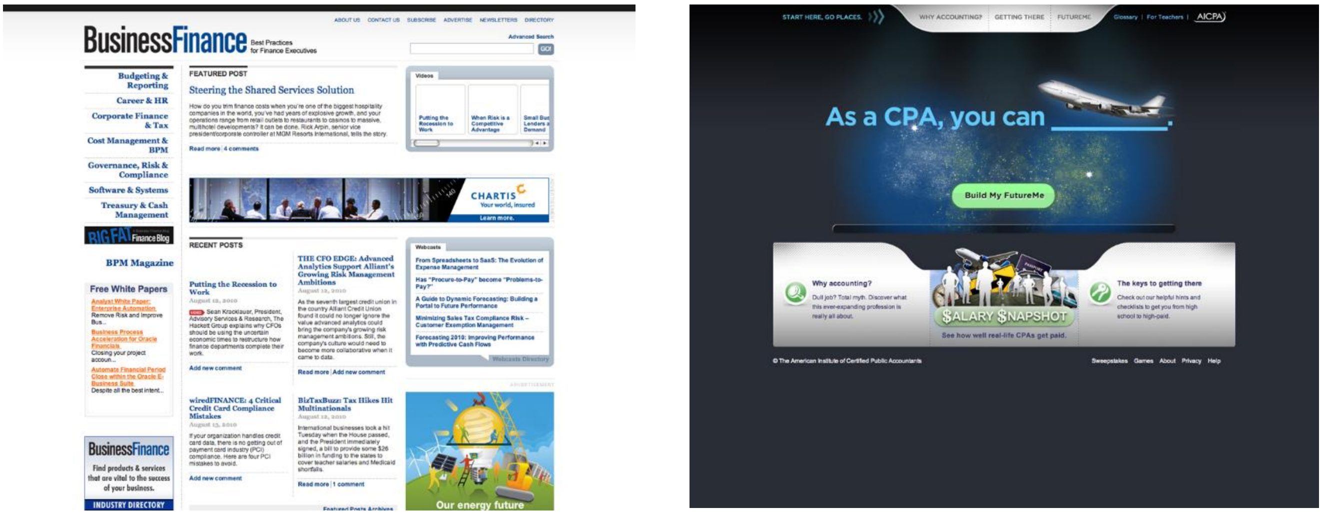
4.5 Conclusion

This study reveals some interesting results which should be kept in mind when creating or optimizing website designs in practice. The time required to make a well-founded first impression visual appeal judgement is incredibly short, as 50ms is already enough to generate this first impression. As mentioned in the introduction, this first impression is critical for making a more thorough opinion and can influence subsequent experiences. Therefore, we can expect a halo effect to occur even within 50ms when visiting websites. Although, it is not proven that this halo effect also leads to higher buying intentions and quality perception (Lorenzo-Romero et al., 2013). Another study has shown that this halo effect of visual appeal is the driver of perceived usability and trustworthiness (Lindgaard et al., 2011). The negative effect of VC suggests that web designers, organizations and other individuals with a website should limit the visual complexity of their websites. To achieve a good first impression, websites should have a high visual appeal and not overload the users with information. Another way to improve website visual appeal is to increase the PT of the websites. Websites should meet the mental model that their visitors have built of (company) websites. Meeting this expectation will enhance the websites PT' and therefore visual appeal is also positively affected.

5. Appendix

5.1 Appendix 1: Visual complexity and prototypicality example

The figure below displays an example of a high VC and high PT website on the left and a low VC and low PT website on the right. Regarding VC it is clear that the left website uses many more elements than the right website, making the overall perceived visual complexity higher (e.g. text, links and images). The PT of the left website is also high, as people find this a typical company website, in contrary to the right website (Tuch et al., 2012). An example for why PT is higher on the left website is for instance the location of the navigation menu and the search bar, which meet the expectations that people have of how a company website should look (Roth et al. 2010). The right website does not provide the user with a search bar and provides only a limited navigation menu on the top of the page, which is atypical. More information regarding the expectations of how company websites should look can be found in the study of Roth et al (2010).



Figuur 15 Example of website characteristics: the left websites is categorized as a high VC and high PT website. The right website is categorized as a low VC and low PT website. Both websites were provided by Tuch et al. (2012)

5.2 Appendix 2: Stimulus selection Tuch et al. (2012)

Tuch et al. (2012) initially selected 464 website homepages which were the most visited websites in the chemicals, energy, accounting, aerospace and defence, automotive, biotechnology and pharmaceuticals and financial services industry. The set of 464 website homepages were reduced towards 270 homepages when the following websites were excluded: 1) websites with no content on the homepage, 2) websites with a shopping basket, 3) websites with advertisements, 4) websites with an archive and finally 5) websites in another language than English or German. These exclusion criteria are based on Roth et al. (2010) which showed that people do not expect shopping baskets, advertisements and archives on company websites. To reduce the set of website homepages even further 267 participant rated the website homepages on visual complexity ("I think this website is of high visual complexity" (Tuch et al., 2010, p.10) and prototypicality ("This website looks like a typical company website" (Tuch et al., 2010, p.10). Every participant rated 30 website homepages, which resulted in at least 14 ratings per website homepage. The ratings between participant on VC and PT were consistent with interclass correlations of .82 for VC and .78 for PT. Finally 120 webpages were selected so a wide range of VC and PT was present in the set of stimuli. In the final set of website homepages VC and PT were independently manipulated as depicted in table 6. Only the main effect was significant and there was no interaction between VC and PT.

	M (SD)	*M (SD)*	*M (SD)*	*F(1-2,114)*	η_p^2
Visual complexity	high	medium	low		
Visual complexity	3.32 (.36)	4.19 (.21)	5.02 (.31)	289.28	.835*
Prototypicality	3.81 (.57)	3.75 (.30)	3.78 (.78)	1.01	.017
Prototypicality	high	-	low		
Visual complexity	4.18 (.29)	-	4.08 (.29)	2.42	.021
Prototypicality	3.78 (.55)	-	4.99 (.30)	215.46	.654*
Visual complexity x prototypicality					
Visual complexity	-	-	-	.28	.005
Prototypicality	-	-	-	.36	.006

Note. * $p < .05$

Table 6 Manipulation stimuli (adapted from Tuch et al., 2010, p.11)

5.3 Appendix 3: Website categories

	LEFT WEBSITE			RIGHT WEBSITE			EXPOSURE DURATION (IN MS)	MIRRORED CATEGORY
CATEGORY	**VC**	**PT**	**Abbr.**	**VC**	**PT**	**Abbr.**		
1	High	Low	VHPL	Low	Low	VLPL	500	
2	High	Low	VHPL	Low	Low	VLPL	50	
3	Low	Low	VLPL	High	Low	VHPL	500	Mirror of 1
4	Low	Low	VLPL	High	Low	VHPL	50	Mirror of 2
5	Low	High	VLPH	Low	Low	VLPL	500	
6	Low	High	VLPH	Low	Low	VLPL	50	
7	Low	Low	VLPL	Low	High	VLPH	500	Mirror of 5
8	Low	Low	VLPL	Low	High	VLPH	50	Mirror of 6
9	High	High	VHPH	Low	Low	VLPL	500	
10	High	High	VHPH	Low	Low	VLPL	50	
11	Low	Low	VLPL	High	High	VHPH	500	Mirror of 9
12	Low	Low	VLPL	High	High	VHPH	50	Mirror of 10
13	Low	High	VLPH	High	Low	VHPL	500	
14	Low	High	VLPH	High	Low	VHPL	50	
15	High	Low	VHPL	Low	High	VLPH	500	Mirror of 13
16	High	Low	VHPL	Low	High	VLPH	50	Mirror of 14
17	High	High	VHPH	High	Low	VHPL	500	
18	High	High	VHPH	High	Low	VHPL	50	
19	High	Low	VHPL	High	High	VHPH	500	Mirror of 17
20	High	Low	VHPL	High	High	VHPH	50	Mirror of 18
21	High	High	VHPH	Low	High	VLPH	500	
22	High	High	VHPH	Low	High	VLPH	50	
23	Low	High	VLPH	High	High	VHPH	500	Mirror of 21
24	Low	High	VLPH	High	High	VHPH	50	Mirror of 22

Table 7 The different website categories including abbreviation, exposure duration and mirror.

5.4 Appendix 4: RMAA results over time

TIMEWINDOW	SPHERICITY CAT	SPHERICITY CAT*TIME	CATEGORY			TIME			CATEGORY X TIME		
VHPL AND VLPH			F-value	Sig.	η_p^2	F-value	Sig.	η_p^2	F-value	Sig.	η_p^2
0 - 20 MS	✓	X	0.862	0.511	0.058	2.806	0.116	0.167	0.711	0.529	0.048
20 - 40 MS	✓	X	1.519	0.195	0.098	0.759	0.398	0.051	0.646	0.606	0.044
40 - 60 MS	✓	X	0.557	0.732	0.038	0.826	0.379	0.056	0.608	0.616	0.042
60 - 80 MS	✓	✓	0.989	0.431	0.066	5.799	**<u>0.030</u>**	0.293	1.116	0.360	0.074
80 - 100 MS	✓	X	3.061	**<u>0.015</u>**	0.179	0.014	0.908	0.001	0.587	0.615	0.040
100 - 120 MS	✓	X	12.451	**<u>0.000</u>**	0.471	0.526	0.480	0.036	0.773	0.486	0.052
120 - 140 MS	✓	✓	17.410	**<u>0.000</u>**	0.554	2.048	0.174	0.128	0.392	0.853	0.027
140 - 160 MS	✓	✓	4.957	**<u>0.001</u>**	0.261	4.039	0.064	0.224	7.233	**<u>0.000</u>**	0.341
160 - 180 MS	✓	✓	2.526	**<u>0.037</u>**	0.153	0.300	0.592	0.021	5.934	**<u>0.000</u>**	0.298
180 - 200 MS	✓	✓	2.492	**<u>0.039</u>**	0.151	0.022	0.883	0.002	3.877	**<u>0.004</u>**	0.217
200 - 220 MS	✓	✓	3.397	**<u>0.008</u>**	0.195	0.704	0.416	0.048	2.719	**<u>0.027</u>**	0.163
220 - 240 MS	✓	✓	4.137	**<u>0.002</u>**	0.228	0.041	0.842	0.003	1.963	0.095	0.123
240 - 260 MS	✓	✓	3.106	**<u>0.014</u>**	0.182	0.657	0.431	0.045	2.546	**<u>0.036</u>**	0.154
260 - 280 MS	✓	X	2.768	**<u>0.024</u>**	0.165	0.140	0.714	0.010	1.118	0.350	0.074
280 - 300 MS	✓	✓	2.652	**<u>0.030</u>**	0.159	0.206	0.657	0.015	0.441	0.819	0.031
300 - 320 MS	✓	✓	1.645	0.160	0.105	0.080	0.781	0.006	0.244	0.941	0.017
320 - 340 MS	✓	✓	0.998	0.425	0.067	0.085	0.774	0.006	0.363	0.872	0.025
340 - 360 MS	✓	✓	1.487	0.205	0.096	0.762	0.397	0.052	0.748	0.590	0.051
360 - 380 MS	✓	X	1.060	0.390	0.070	0.328	0.576	0.023	0.885	0.446	0.059
380 - 400 MS	✓	✓	0.866	0.373	0.026	0.061	0.809	0.004	0.194	0.964	0.014
400 - 420 MS	✓	✓	0.659	0.656	0.045	0.386	0.544	0.027	0.607	0.695	0.042
420 - 440 MS	✓	✓	1.510	0.198	0.097	0.423	0.526	0.029	0.764	0.579	0.052
440 - 460 MS	✓	✓	2.560	**<u>0.035</u>**	0.155	1.314	0.271	0.086	0.758	0.583	0.051
460 - 480 MS	✓	✓	2.964	**<u>0.017</u>**	0.175	0.602	0.451	0.041	0.638	0.672	0.044
480 - 500 MS	✓	✓	2.227	0.061	0.137	0.840	0.375	0.057	0.213	0.956	0.015

Table 8 Two-way RMAA analyses 1 results with constant category VHPL and VLPH. Note that it is indicated whether the assumption of sphericity is met.

When this assumption is not met, the Greenhouse-Geisser correction is used. Also, all significant results with an alpha of 0.045 are bold and underlined

TIMEWINDOW	SPHERICITY CAT	SPHERICITY CAT*TIME	CATEGORY			TIME			CATEGORY X TIME		
VLPL AND VHPH			F-value	Sig.	η_p^2	F-value	Sig.	η_p^2	F-value	Sig.	η_p^2
0 - 20 MS	✓	X	0.495	0.779	0.034	0.319	0.581	0.022	1.164	0.330	0.077
20 - 40 MS	✓	X	0.426	0.691	0.030	0.320	0.581	0.022	0.717	0.552	0.049
40 - 60 MS	✓	X	0.375	0.864	0.026	0.438	0.519	0.030	0.720	0.543	0.049
60 - 80 MS	✓	✓	1.136	0.350	0.075	0.029	0.866	0.002	2.142	0.070	0.133
80 - 100 MS	✓	X	0.196	0.963	0.014	0.166	0.689	0.012	0.450	0.726	0.031
100 - 120 MS	✓	X	3.183	**0.012**	0.185	0.018	0.894	0.001	0.841	0.465	0.057
120 - 140 MS	✓	✓	5.541	**0.002**	0.284	0.145	0.709	0.010	0.816	0.542	0.055
140 - 160 MS	✓	✓	1.781	0.128	0.113	2.939	0.108	0.174	7.114	**0.000**	0.337
160 - 180 MS	✓	✓	3.104	**0.014**	0.181	6.098	**0.027**	0.303	4.801	**0.001**	0.255
180 - 200 MS	✓	✓	2.769	**0.024**	0.165	5.894	**0.029**	0.296	2.935	**0.018**	0.173
200 - 220 MS	✓	✓	1.854	0.114	0.177	4.986	**0.042**	0.263	1.957	0.096	0.123
220 - 240 MS	✓	✓	2.480	**0.040**	0.150	2.875	0.112	0.170	1.231	0.304	0.081
240 - 260 MS	✓	✓	1.755	0.134	0.111	0.962	0.343	0.064	2.523	**0.037**	0.153
260 - 280 MS	✓	X	1.773	0.130	0.112	0.134	0.720	0.009	1.052	0.382	0.070
280 - 300 MS	✓	✓	2.097	0.076	0.130	0.020	0.890	0.001	0.472	0.796	0.030
300 - 320 MS	✓	✓	1.598	0.172	0.102	0.030	0.865	0.002	0.224	0.951	0.016
320 - 340 MS	✓	✓	1.081	0.379	0.072	0.185	0.673	0.013	0.373	0.866	0.026
340 - 360 MS	✓	✓	1.167	0.334	0.077	0.006	0.939	0.000	1.173	0.331	0.077
360 - 380 MS	✓	✓	1.274	0.285	0.083	0.120	0.734	0.008	0.898	0.488	0.060
380 - 400 MS	✓	✓	0.514	0.764	0.035	0.014	0.906	0.001	0.199	0.962	0.014
400 - 420 MS	✓	✓	0.500	0.775	0.034	0.008	0.928	0.001	0.728	0.605	0.049
420 - 440 MS	✓	✓	1.490	0.204	0.096	0.037	0.851	0.003	0.937	0.463	0.063
440 - 460 MS	✓	✓	2.058	0.081	0.128	0.270	0.611	0.019	1.014	0.416	0.068
460 - 480 MS	✓	✓	2.797	**0.023**	0.167	0.060	0.810	0.004	0.816	0.542	0.055
480 - 500 MS	✓	✓	1.882	0.109	0.118	0.060	0.810	0.004	0.337	0.889	0.024

Table 9 Two-way RMAA 2 results with constant category VLPL and VHPH. Note that it is indicated whether the assumption of sphericity is met. When this assumption is not met, the Greenhouse-Geisser correction is used. Also, all significant results with an alpha of 0.045 are bold and underlined

7.5 Appendix 5: Effect of VC and PT on PO7

VHPL	HIGH VC				HIGH PT			
	P	M	SE	98.75% CI	P	M	SE	98.75% CI
80 - 100 MS	0.515	0.241	0.360	[-0.792, 1.273]	0.065	0.645	0.322	[-0.277, 1.566]
100 - 120 MS	0.025	0.996	0.396	[-0.137, 2.129]	0.032	0.876	0.369	[-0.180, 1.932]
120 - 140 MS	**0.001**	**0.863**	**0.209**	**[0.265, 1.462]**	**0.541**	**0.294**	**0.468**	**[-1.048, 1.635]**

Table 10 Pairwise comparisons to analyse the effect of VC and PT when VHPL is constant.

VLPH	HIGH VC				HIGH PT			
	P	M	SE	98.75% CI	P	M	SE	98.75% CI
80 - 100 MS	0.251	0.374	0.313	[-0.521, 1.270]	0.078	0.689	0.363	[-0.350, 1.729]
100 - 120 MS	0.006	1.125	0.344	[0.140, 2.109]	0.16	1.092	0.401	[-0.057, 2.240]
120 - 140 MS	**0.001**	**1.798**	**0.432**	**[0.561, 3.035]**	**0.006**	**1.308**	**0.406**	**[0.146, 2.471]**

Table 11 Pairwise comparisons to analyse the effect of VC and PT when VLPH is constant.

VLPL	HIGH VC				HIGH PT			
	P	M	SE	98.75% CI	P	M	SE	98.75% CI
100 - 120 MS	0.745	0.145	0.439	[-1.110, 1.400]	0.046	0.961	0.439	[-0.295, 2.217]
120 - 140 MS	**0.850**	**0.065**	**0.338**	**[-0.903, 1.033]**	**0.007**	**1.535**	**0.492**	**[0.127, 2.943]**

Table 12 Pairwise comparisons to analyse the effect of VC and PT when VLPL is constant.

VHPH	HIGH VC				HIGH PT			
	P	M	SE	98.75% CI	P	M	SE	98.75% CI
100 - 120 MS	0.586	0.274	0.492	[-1.134, 1.682]	0.043	0.745	0.334	[-0.212, 1.703]
120 - 140 MS	0.053	1.000	0.473	[-0.356, 2.356]	0.351	0.521	0.539	[-1.024, 2.066]

Table 13 Pairwise comparisons to analyse the effect of VC and PT when VHPH is constant.

6. References

Annett, M. (1970). A classification of hand preference by association analysis. *British journal of psychology*, 61(3), 303-321.

Bach, M. (2006). The Freiburg Visual Acuity Test-variability unchanged by post-hoc re-analysis. *Graefe's Archive for Clinical and Experimental Ophthalmology*, 245(7), 965-971.

Bar, M., Neta, M., & Linz, H. (2006). Very first impressions. *Emotion*, *6*(2), 269.

Breitmeyer, B., & Öğmen, H. (2006). *Visual masking: Time slices through conscious and unconscious vision* (No. 41). Oxford University Press.

Celesia, G. G., & Brigell, M. (1992). Event-related potentials. *Current Opinion in Neurology*, *5*(5), 733-739.

Eimer, M. (1996). The N2pc component as an indicator of attentional selectivity. *Electroencephalography and clinical neurophysiology*, *99*(3), 225-234.

Enns, J. T., & Di Lollo, V. (2000). What's new in visual masking? *Trends in cognitive sciences, 4*(9), 345-352.

Foxe, J. J., Simpson, G. V., & Ahlfors, S. P. (1998). Parieto-occipital~ 1 0Hz activity reflects anticipatory state of visual attention mechanisms. *Neuroreport*, *9*(17), 3929-3933.

Goldstein, E.B., 2009. *Sensation and Perception*. 8th ed. Belmont: Wadsworth Publishing.

Ishihara, S. (1976). *Test for colour blindness, 38 plates edition*. Tokyo: Kanehara Shuppen Co.

Leder, H., Belke, B., Oeberst, A., & Augustin, D. (2004). A model of aesthetic appreciation and aesthetic judgments. *British journal of psychology*, 95(4), 489-508.

Leuthesser, L., Kohli, C. S., & Harich, K. R. (1995). Brand equity: the halo effect measure. *European Journal of Marketing, 29*(4), 57-66.

Lindgaard, G., Dudek, C., Sen, D., Sumegi, L., & Noonan, P. (2011). An exploration of relations between visual appeal, trustworthiness and perceived usability of homepages. *ACM Transactions on Computer-Human Interaction (TOCHI)*, *18*(1), 1.

Lindgaard, G., Fernandes, G., Dudek, C., & Brown, J. (2006). Attention web designers: You have 50 milliseconds to make a good first impression!. *Behaviour & information technology*, *25*(2), 115-126.

Lindgaard, G., Litwinska, J., & Dudek, C. (2008). Judging web page visual appeal: Do east and west really differ. *Proc. IADIS, 8*, 157-164.

Lorenzo-Romero, C., Constantinides, E., María-del-Carmen & Alarcón-del-Amo (2013). Web aesthetics effects on user decisions: impact of exposure length on website quality perceptions and buying intentions. *Journal of Internet Commerce*, *12*(1), 76-105.

Luck, S. J. (2012). Event-related potentials. *APA Handbook of Research Methods in Psychology: Volume 1, Foundations, Planning, Measures, and Psychometrics.*

Maner, J. K., Kenrick, D. T., Becker, D. V., Delton, A. W., Hofer, B., Wilbur, C. J., & Neuberg, S. L. (2003). Sexually selective cognition: beauty captures the mind of the beholder. *Journal of personality and social psychology*, *85*(6), 1107.

Nevid, J. S. (2012). *Psychology: concepts and applications*. Cengage Learning.

Öhman, A., Flykt, A., & Esteves, F. (2001). Emotion drives attention: detecting the snake in the grass. *Journal of experimental psychology: general*, *130*(3), 466.

Reber, R., Schwarz, N., & Winkielman, P. (2004). Processing fluency and aesthetic pleasure: Is beauty in the perceiver's processing experience? *Personality and social psychology review*, 8(4), 364-382.

Robins, D., & Holmes, J. (2008). Aesthetics and credibility in web site design. *Information Processing & Management*, *44*(1), 386-399.

Roth, S. P., Schmutz, P., Pauwels, S. L., Bargas-Avila, J. A., & Opwis, K. (2010). Mental models for web objects: Where do users expect to find the most frequent objects in online shops, news portals, and company web pages? *Interacting with computers*, 22(2), 140-152.

Schoormans, J. P., & Robben, H. S. (1997). The effect of new package design on product attention, categorization and evaluation. *Journal of Economic Psychology*, *18*(2), 271-287.

Seidman, G., & Miller, O. S. (2013). Effects of gender and physical attractiveness on visual attention to Facebook profiles. *Cyberpsychology, Behavior, and Social Networking*, *16*(1), 20-24.

Thielsch, M. T., & Hirschfeld, G. (2012). Spatial frequencies in aesthetic website evaluations–explaining how ultra-rapid evaluations are formed. *Ergonomics*, 55(7), 731-742.

Thut, G., Nietzel, A., Brandt, S. A., & Pascual-Leone, A. (2006). α-Band electroencephalographic activity over occipital cortex indexes visuospatial attention bias and predicts visual target detection. *Journal of Neuroscience*, *26*(37), 9494-9502.

Tractinsky, N., Cokhavi, A., Kirschenbaum, M., & Sharfi, T. (2006). Evaluating the consistency of immediate aesthetic perceptions of web pages. *International journal of human-computer studies*, *64*(11), 1071-1083.

Tuch, A. N., Bargas-Avila, J. A., Opwis, K., & Wilhelm, F. H. (2009). Visual complexity of websites: Effects on users' experience, physiology, performance, and memory. *International journal of human-computer studies*, *67*(9), 703-715.

Tuch, A., Kreibig, S., Roth, S., Bargas-Avila, J., Opwis, K., & Wilhelm, F. (2011). The role of visual complexity in affective reactions to webpages: subjective, eye movement, and cardiovascular responses. *IEEE transactions on affective computing*, *2*(4), 230-236.

Tuch, A. N., Presslaber, E. E., StöCklin, M., Opwis, K., & Bargas-Avila, J. A. (2012). The role of visual complexity and prototypicality regarding first impression of websites: Working towards understanding aesthetic judgments. *International Journal of Human-Computer Studies*, *70*(11), 794-811.

Van der Lubbe, R. H., Bundt, C., & Abrahamse, E. L. (2014). Internal and external spatial attention examined with lateralized EEG power spectra. *Brain research*, 1583, 179-192.

Van der Lubbe, R. H., Jaśkowski, P., Wauschkuhn, B., & Verleger, R. (2001). Influence of time pressure in a simple response task, a choice-by-location task, and the Simon task. *Journal of Psychophysiology*, *15*(4), 241.

Van der Lubbe, R. H., Szumska, I., & Fajkowska, M. (2016). Two sides of the same coin: ERP and wavelet analyses of visual potentials evoked and induced by task-relevant faces. *Advances in cognitive psychology*, *12*(4), 154.

Van der Lubbe, R. H., & Utzerath, C. (2013). Lateralized power spectra of the EEG as an index of visuospatial attention. *Advances in cognitive psychology*, *9*(4), 184.

Van Hooff, J. C., Crawford, H., & Van Vugt, M. (2010). The wandering mind of men: ERP evidence for gender differences in attention bias towards attractive opposite sex faces. *Social cognitive and affective neuroscience*, *6*(4), 477-485.

Van Schaik, P., & Ling, J. (2009). The role of context in perceptions of the aesthetics of web pages over time. *International Journal of Human-Computer Studies*, *67*(1), 79-89.

Wascher, E., & Wauschkuhn, B. (1996). The interaction of stimulus-and response-related processes measured by event-related lateralizations of the EEG. *Electroencephalography and clinical neurophysiology*, *99*(2), 149-162.

Willis, J., & Todorov, A. (2006). First impressions making up your mind after a 100-ms exposure to a face. *Psychological science*, *17*(7), 592-598.

Winkielman, P., Halberstadt, J., Fazendeiro, T., & Catty, S. (2006). Prototypes are attractive because they are easy on the mind. *Psychological science*, *17*(9), 799-806.

Xing, J., Manning, C., 2005. Complexity and automation displays of air traffic control: Literature review and analysis. *Tech. rep., U.S. Department of Transportatioin*, Office of Aerospace Medicine.