

Electrophysiological correlates of aesthetic processing of webpages: A comparison of experts and  
laypersons

Jens Bölte<sup>1</sup>, Thomas Hösker<sup>1</sup>, Gerrit Hirschfeld<sup>2</sup>, Meinald Thielsch<sup>1</sup>

<sup>1</sup>Westfälische Wilhelms-Universität Münster, Münster, Germany

<sup>2</sup>University of Applied Sciences Osnabrück, Germany

Corresponding author:

Jens Bölte

Institut für Psychologie

Westfälische Wilhelms-Universität Münster

Fliedner Str. 21

48159 Münster, Germany

**Abstract**

We investigated whether design experts or laypersons evaluate webpages differently. Twenty participants, 10 experts and 10 laypersons, judged the aesthetic value of a webpage in an EEG-experiment. Screenshots of 150 webpages, judged aesthetic or unaesthetic by different sample of 136 participants, served as stimulus material. Behaviorally experts and laypersons evaluated unaesthetic webpages similarly but they differed in their evaluation of aesthetic webpages. Experts evaluated aesthetic webpages more often as unaesthetic than laypersons did. The ERP-data show difference between experts and laypersons and between aesthetic and unaesthetic webpages. In a time-window of 110-130ms after stimulus onset, aesthetic webpages elicited a more negative EEG-amplitude than unaesthetic webpages. Experts had more negative EEG-amplitude than layperson. There was no interaction of expertise and aesthetics. This patterning of results continued until a time window of 600-800ms in which group and aesthetic differences diminished. An interaction of perceiver and object, as interactionist theories postulate, was absent in the EEG-data. Experts seem to process the stimuli in a more thorough manner than laypersons. The early activation differences between aesthetic and unaesthetic webpages is in contrast with some theories of aesthetic processing and has not been reported before.

## Introduction

Humans appreciate aesthetic entities in various contexts: Sometimes entities are created for aesthetic and intellectual purposes only as in fine arts. Sometimes, customers should be attracted by the aesthetic properties of a product. Humans also evaluate the aesthetic properties of everyday objects such as tableware, TV-sets, or food. Aesthetic judgments play a role in many aspects of human life (Hoyer & Stokburger-Sauer, 2012) differing between cultures, situations, educational background, expertise and other individual properties (Jacobsen, 2006).

We focus in our research on the differences that experts and laypersons might show in evaluating the aesthetic quality of webpages. Experts and laypersons differ in their aesthetic judgments in various domains such as art, music, movies, software code or even facades of houses (Hasse & Weber, 2012; Kozbelt, Dexter, Dolese & Seidel, 2012; Müller, Höfel, Brattico & Jacobsen, 2009; Silvia, 2013; Silvia & Berg, 2011). Webpages are a modern medium of communication serving nearly every aspect of human living. They are evaluated not only in terms of functionality but also in terms of their aesthetic properties (Lavie & Tractinsky, 2004; Moshagen & Thielsch, 2010). While electrophysiological responses to music, for instance, have been described (Müller et al., 2009; Müller, Höfel, Brattico & Jacobsen, 2010) it is unclear whether similar event-related potentials will show up when evaluating webpages.

**Comment [A1]:** Will be present-  
A very minor point

## Aesthetics of webpages

Aesthetics has become a core construct in research on human computer interaction and especially webpage perception (Moshagen & Thielsch, 2010). Similar to other everyday stimuli such as logos (Handy et al., 2010), webpages (Lindgaard et al., 2006; Thielsch & Hirschfeld, 2012; Tractinsky, Cokhavi, Kirschenbaum & Sharfi, 2006) are spontaneously processed at an aesthetic level (for an overview see Tuch, Presslauer, et al., 2012). Webpage aesthetics, i.e. the

**Comment [A2]:** Perception?  
Do you mean understanding or perhaps perception and cognition of webpages – perception implies interest in low level processing

“immediate pleasurable subjective experience that is directed toward an object and not mediated by intervening reasoning” (Moshagen & Thielsch, 2010, p. 690), has an impact on various constructs: for example, perceived usability, credibility, satisfaction, preference, urge to buy impulsively or intention to revisit (for overviews see Lee & Koubek, 2012; Moshagen & Thielsch 2010; Tuch, Roth, Hornbæk, Opwis & Bargas-Avila, 2012b).

The current research suggests that aesthetic responses to webpages occur immediately at first sight (e.g., Lindgaard et al., 2006 & 2011; Thielsch & Hirschfeld, 2012; Tractinsky et al., 2006). This is not only in line with results of prior research on the aesthetics of art, but as well of high practical relevance, as users’ first impressions are very relevant for the decision of whether a particular webpage is explored deeper or another one is searched for (cf. Thielsch, Blotenberg & Jaron, 2014). However, the major purpose of a webpage is not an aesthetic one different than art. Webpages are mostly designed to provide some kind of information, mostly in an interactive manner (cf. ISO, 2006; Thielsch, et al., 2014).

**Comment [A3]:** Sentence does not make sense

### Models of aesthetic processing

Fechner (1876) pioneered the empirical investigation of aesthetic processing. A field that developed to the field of neuroaesthetics (Zeki, 2001) nowadays. In some theories, mostly earlier ones, it is assumed that object properties determine the aesthetic evaluation. Properties such as balance and proportion (Arnheim, 1974; Birkhoff, 1933; Fechner, 1876), novelty and prototypicality (Hekkert, Snelders & van Wieringen, 2003; Hekkert & van Wieringen, 1990), contrast and clarity (Gombrich, 1995; Solso, 2003) as well as the relation between object features, particularly between simplicity and complexity (Birkhoff, 1933; von Ehrenfels, 1890; Eysenck, 1941) are the basis of the aesthetic evaluation. In recent years, interactionist perspectives on aesthetics that focus on the interplay between the observer and the object have

**Comment [A4]:** sentence

been put forward (e.g., Berlyne, 1971; Chatterjee, 2004; Jacobsen, 2006; 2010; Leder, Belke, Oeberst & Augustin, 2004; Leder & Nadal, 2014; Reber, Schwarz & Winkielman, 2004). The idea that aesthetic processing is best described in terms of distinct processing stages form the basis of such perspective.

According to the framework by Leder et al. (2004; Leder & Nadal, 2014), aesthetic processing follows five stages: perception, implicit classification, explicit classification, cognitive mastering and evaluation. Cognitive processing is accompanied by affective states in an interactive manner. There are two outcomes: First, the aesthetic judgment claiming to what degree the object in question meets the normative standards of aesthetic. Second, the aesthetic emotion or appraisal is characterized by the affective experiences during the aesthetic processing.

While this framework of aesthetic appreciation and aesthetic judgments (Leder et al., 2004) emphasizes psychological processes, the model of visual neuroaesthetics by Chatterjee (2004) focuses on the neuronal basics of aesthetic processing. In this model, early and intermediate visual processing steps are similar for aesthetic as well as non-aesthetic processing. During early vision, basic stimulus parameters are processed such as color, shape, and contrast (Marr, 1982). Thereafter, the stimulus parameters are grouped together during intermediate vision. In the latest phase, object recognition and affective and aesthetics evaluation takes place. Aesthetically relevant object features trigger attention processes during early and intermediate stages of visual processing. Thereby, the processing of the aesthetic characteristics is enhanced in later phases of visual processing in terms of a feed forward system (Chatterjee, 2004).

## EEG studies on aesthetic processing

EEG studies are a valuable tool to shed some light on the timing of aesthetic processing.

The comparison of aesthetic and descriptive assessments of geometric patterns showed that aesthetic processing is a two-step process characterized by an early frontal negative potential (ERAN) between 300 and 400 ms after stimulus onset that occurs for unaesthetic judgments only.

The second step is associated to a late positive potential (LPP) in central and parietal electrodes between 440 and 880 ms peaking around 600 ms after stimulus onset. ERAN supposedly reflects impression formation while the LPP indicates evaluative categorization of the stimulus (Höfel & Jacobsen, 2007a; Höfel & Jacobsen, 2007b; Jacobsen & Höfel, 2003).

Using an oddball-paradigm, Wang, Huang, Ma and Li (2012) eliminated an instructed aesthetic evaluation by the participants. They observed an enhanced P2 for less beautiful pendants in comparison to more beautiful pendants. A P2 is taken as an indicator of early affective evaluation, for instance of words or pictures that elicit negative feelings (Huang & Luo, 2006). Therefore, Wang et al. concluded that the P2 they observed reflects early automatic emotional processes that accompany aesthetic processing. Alternatively, Carretie, Mercado, Tapia and Hinojosa (2001) suggested that the P2 reflects attentional processes. Given that Wang et al. used an oddball-paradigm that is suited to investigate automatic attentional processes this alternative interpretation cannot be ruled out.

De Tommaso et al. (2008) asked their participants to categorize targets as beautiful, neutral or ugly among standard stimuli (green panel). Targets were either famous paintings or geometric shapes. Target stimuli resulted in a larger N2-amplitude than standard stimuli and a larger P3 (Experiment 1). A larger P300 was observed for geometric shapes than for paintings. Neutral paintings elicited a larger N2 than did beautiful paintings. In Experiment 2, participants

**Comment [A5]:** you mention early negativity enhanced for unaesthetic judgments (as compared to ??? clarify) also – you do not outline the conditions that modulate the LPP, does this mean you are stating it is the same as the earlier negative component? I.e. larger LPP for unaesthetic compared with? Please clarify

**Comment [A6]:** for unaesthetic only? clarify

**Comment [A7]:** Could you please clarify what is meant here - how did the authors eliminate an evaluation? Do you mean the LPP was reduced? Please clarify

**Comment [A8]:** Is this a posterior or anterior/central P2 please give full information of the component (what is the time window and spatial distribution)

**Comment [A9]:** Is this a posterior P2? Comparing early ERPs to 2 stimulus categories will lead to differences due to low level features in stimuli (spatial frequency of images) P2 involved in stimulus categorization certainly

**Comment [A10]:** N2a or N2b? I.e. N2 anterior or N2 posterior these two components have different functional characteristics –please make it clear whether you are referring to the anterior or posterior N2

had to indicate the presence of a target. Here, N2-amplitudes was again larger for targets than for standards but there was no modulation due to aesthetic value. However, beautiful target stimuli elicited a larger P3 than ugly or neutral target stimuli, all of which elicit a larger P3 than standard stimuli. The authors argue that task differences bring about the different results. The aesthetics categorization task in Experiment 1 is more difficult for an aesthetically indifferent target than for a beautiful or ugly target. The P3-increase for geometric shapes compared to paintings reflects the more difficult aesthetic categorization process for shapes than for paintings. Such differences were absent in Experiment 2, suggesting that attention levels as measured by the P3 are not influenced by stimulus type but only by aesthetic value.

**Comment [A11]:** This section needs to be made a little clearer

### Individual differences: Experts versus laypersons

These empirical findings reveal the impact of stimulus properties on the different suggested processing steps (Chatterjee, 2004; Jacobsen, 2006; Leder et al., 2004). However, varying subject characteristics such as age (Thielsch, 2008), gender (Cela-Conde et al., 2009; Tuch, Bargas-Avila & Opwis, 2010) or domain specific expertise that could influence aesthetic processing were not taken into account (Chevalier & Ivory, 2003; Park, Choi & Kim, 2004). Especially, the model by Leder et al. that refers to implicit or explicit knowledge in long-term memory allows predicting differences in processing by experts and laypersons.

In the context of face perception research, it has been shown that expertise for face-like objects can be acquired by training (e.g., Gauthier & Tarr, 1997) and that the processing of faces or other well-trained materials goes along with distinct brain activity (Bentin, Allison, Puce, Perez & McCarthy, 1996; Gauthier, Skudlarski, Gore & Anderson, 2000; Tanaka & Curran, 2001). Thus, expertise for specific objects can be acquired and is reflected in distinguishable cortical activity (Scott, Tanaka, Sheinberg & Curran, 2006).

**Comment [A12]:** Tanaka refers to specific ERP components as neural index of visual expertise (for example in explicit recognition in experts N250) Why do you give neuroimaging examples when this is ERP paper? Perhaps neuroimaging evidence is fine but done leave out the ERP detail/evidence in favour of the neuroimaging detail- this is what is relevant

For example, architects show higher hippocampus, precuneus, orbitofrontal cortex and gyrus cinguli activity than laypersons when asked to rate the aesthetics of buildings (Kirk, Skov, Christensen & Nygaard, 2009). Moreover, designer and laypersons differ in their cortical activity during a design task (Kowatari et al., 2009). Similarly, art specific expertise is accompanied by functional and structural modifications, i. e. higher cortical activation during color processing and higher density of grey matter in area V4 (Long, Peng, Chen, Jin & Yao, 2011). Furthermore, laypersons and experts differ in perceptual exploration (Hekkert & van Wieringen, 1996), processing of complexity (Reber et al., 2004) and aesthetic preferences (Hekkert & van Wieringen, 1996). Generally, expertise is presumably characterized by stronger than usual neuronal connections of specific representations or processes which result in high accessibility of these representations (Cheung & Bar, 2012; Harel, Gilaie-Dotan, Malach & Bentin, 2010).

Müller et al. (2010) investigated the differential processing of short piano sequences by laypersons and experts using EEG. They observed a larger P2 amplitude for experts than for laypersons and a larger ERAN, which peaked 200 ms after stimulus onset. Müller et al. assumed that the enhanced P2 reflects extended neural representations for musical stimuli in experts. The larger ERAN indicates a more thorough processing of the stimuli by the experts than by the laypersons because it was only observed with mild harmonic violations.

**Comment [A13]:** Is this P2 the opposite dipole of the anterior ERAN- they seem to be in the same time window? Please notify the reader if this is posterior P2

Using paintings, filtered copies of these paintings and plain-color stimuli as visual stimuli, Pang, Nadal, Müller-Paul, Rosenberg and Klein (2013) could show that paintings elicit larger P3b components than their filtered copies or plain-color stimuli. Experts, determined by a questionnaire, had smaller EEG-amplitudes than laypersons. Pang et al. argued that this reduced activity is a consequence of neural efficiency due to increased practice. Clearly, this observation contradicts the one reported by Müller et al. (2010). Müller et al. observed increased neuronal activation for experts in comparison to laypersons. However, various differences between the

**Comment [A14]:** P3 is not just related to stimulus but also interacts with task- what task did the authors use

**Comment [A15]:** P3b?

**Comment [A16]:** Please mention the participants task given this is a P3 result (as comment above)



experiments, visual vs. auditory stimulation, passive viewing vs. explicit judgment, prohibit firm conclusions. The idea that expertise is realized by stronger neuronal connections is compatible with both patterns (Cheung & Bar, 2012; Harel et al., 2010). Taken together the research has revealed several differences between experts and laypersons when it comes to the processing of aesthetic stimuli. It is unknown whether these differences map to the processing stages surmised by current models of aesthetic perception (Chatterjee, 2004; Jacobsen, 2006; Leder et al., 2004). In addition, it is still undetermined whether expertise results regularly in more or less neuronal activity.

**Comment [A17]:** And could simply reflect the use of different reference electrodes across experiments along with task related effects

**Comment [A18]:** I'm not sure what you mean by this please clarify- stronger neuronal connections is not the same as enhanced P3? Are you referring to connectivity between distinct brain regions?

**Comment [A19]:** Does 'both patterns' here mean- increases and reductions in P3b amplitude Please clarify?

## Research question

The research presented above shows that experts and laypersons differ in the processing of stimuli, including the evaluation of aesthetic stimulus qualities. Most often music or stimuli constructed for the particular research question were used for such research questions. We want to extend this research by using commonly experienced stimuli such as webpages.

**Comment [A20]:** What does this mean? Reductions or enhancements in ERP components will depend on the particular component and importantly the reference electrode used in a particular study. Certainly more or less 'amplitude' does not equate with more or less effort?

Can you please clarify/simplify this statement  
Question does expertise modulate ERP waveforms and if so are they early perceptual or later cognitive (EDAN or LPP)  
I think this is the question you are asking- one with regard to the time course of expert aesthetic processing?

**Comment [A21]:**

**Comment [A22]:** Your analyses proceeds with a number of time windows during early and late components- do you have any rationale for analyzing all these time windows?

**Comment [A23]:** I think you need to make a clearer hypothesis-/prediction as to which ERP components you expect to see effects of expertise and effects of aesthetic/unaesthetic

## Materials and Methods

### Participants

There were 20 participants, 10 were experts (5 females, mean age: 32.2 years,  $SD = 12.5$  years) and 10 laypersons (5 females, mean age: 31 years,  $SD = 11.4$  years). Experts were either professionals in or students of design, graphic design or digital media design. They had a mean experience in the area of 10.8 years ( $SD = 10.7$  years). Experts attach more value to the visual-aesthetic product design than laypersons in the Centrality of Visual Product Aesthetics questionnaire (CVPA; Bloch, Brunel & Arnold, 2003; German version Thielsch, 2008; experts:

**Comment [A24]:** It would be useful to know the range of the two groups (youngest-oldest) just because of sensitivity of EEG to changes across the lifespan – but this is a minor point (just for completeness)

$M_{CVPA} = 4.18$ ,  $SD = .41$ ; laypersons:  $M_{CVPA} = 2.9$ ,  $SD = .52$ ;  $t(18) = 6.08$ ,  $p < .001$ ,  $d = 1.6$ ).

Laypersons were matched to experts in terms of gender, age and education.

Participants received 10.00 € for their participation. They had unimpaired or corrected-to-normal visual acuity as well as normal color perception. Our study did not require the approval of our local Ethics Committee at the Department of Psychology at the University of Münster as we performed a non-clinical website evaluation study and only non-invasive measures (Ratings, reaction times, EEG). The task was to assess different websites in respect to their aesthetics. No treatments or false feedbacks were given; no potential harmful evaluation methods were used. Participation was completely voluntary and participants could drop out at any time without any negative consequences. Informed consent was obtained from all participants. All data were stored only using an anonymous ID for each participant.

**Comment [A25]:** Is this standard? Ethics are usually gained in university/research settings when human subjects are involved?? Does the journal/publisher require ethic to be granted by a panel

**Comment [A26]:** Remove 'completely' as not strictly true- stated participants are paid volunteers

## Materials

We used 150 webpage-screenshots in the experiments. These screenshots were selected from a larger set of 300 webpage-screenshots that had been rated in terms of their aesthetic attractiveness by another 136 participants (110 females, mean age: 25.2 years,  $SD = 6.77$  years) using the short version of the Visual Aesthetics of Websites Inventory (VisAWI-S, Moshagen & Thielsch, 2013). Websites known by more than 25% of the participants were excluded from further analysis. Based on the results, two sets of 75 webpage-screenshots were created differing significantly in aesthetic assessment ( $M_{\text{aesthetic}} = 5.11$ ,  $SD = .39$ ;  $M_{\text{unaesthetic}} = 3.27$ ,  $SD = .45$ ;  $d = 5.76$ ;  $\chi^2 = 146.03$ ,  $p < .001$ ). The webpage-screenshots come from ten different content domains (download & software, e-commerce, e-learning, entertainment, e-recruiting, information, corporate webpages, social software, search engines, and web portals). Information about the

browser was removed using Adobe Photoshop, Version CS5.1 ©. Webpages were 1280 x 780 to 1264 x 765 pixels large (Welch  $t(146.61) = -1.31, p = .192$ , unaesthetic mean: 1012742 pixel, SD: 5820, aesthetic mean: 1013931 pixel, SD: 5278). The average luminance per pixel relative to white (white having a value of 1.0, black having a value of 0) did not differ from each other (Welch  $t(142.63) = 1.677, p = .095$ , unaesthetic mean: .785, SD: 0.149, aesthetic: .739, SD: 0.182). Also, unaesthetic and aesthetic webpages did not differ in contrast (Welch  $t(147.93, p = .408$ ; unaesthetic: .261, SD: 0.069, aesthetic: .252, SD: 0.068). Aesthetic and unaesthetic webpages differed in complexity measured in byte (Welch  $t(136.55) = 3.59, p < .001$ , unaesthetic mean: 681 kb, SD: 199; aesthetic: mean: 578 kb, SD: 147). The later difference might reflect design differences between unaesthetic (e.g. cluttered layout) and aesthetic webpages (e.g., clear, structured layout).<sup>1</sup>

**Comment [A27]:** It might be useful to simply have the visual angle for the two different stimulus types- When accounting for viewing distance and size on screen I doubt there any much difference It is usual to report visual angle for ERP study

## Apparatus

The experiment was controlled by Presentation Version 16.04.25.12 (Neurobehavioral Systems, Albany, CA, USA). Stimuli were presented on a Samsung Sync-Master 2233, 1680 x 1050 pixels, screen refresh rate 120 Hz. Responses were collected using a Cedrus response-pad RB 830. The EEG was digitized with a sampling frequency of 256 Hz using 32 sintered Ag/AgCL-electrodes in 10-20 system (Advanced Neuro Technology, Enschede, The Netherlands). We used an online low-pass half-power filter of 69.12 Hz and an average reference for recording. Impedance was kept below 5 kOhm. Vertical EOG was measured by placing a bipolar electrode beneath and above the left eye. Horizontal EOG was measured by placing a bipolar electrode at the outer canthus of each eye. AFz was used as ground electrode.

**Comment [A28]:** Yes, there probably are differences in spatial frequency between the two stimulus types – this probably does interact with aesthetics and is kind of given in the research question but does not detract from 'interestingness' but needs to be taken into account in interpretation of ERP data (in particular early sensory components which are of course sensitive to spatial frequency)

**Comment [A29]:** I think it is important to give the electrode montage you used- simply because you mention on the next page that you re-referenced to linked mastoid- and you have not outlined in this section that you recorded mastoid electrodes so that means the reader does not know if you simply recorded 30 electrodes on the scalp plus 2 on the mastoid (which of course would be too low a resolution/montage for an average online reference recording! So please report the electrodes you actually record from (all of them) and specify whether the montage is in accord the 10-20 system layout as is typical for this kind of work!

## Procedure

All participants were interviewed via a telephone-interview in which demographic information, handedness, visual acuity, neurological or psychiatric disorders were assessed. Upon arrival in laboratory, participants were informed about the course of the experiment. They first completed the Centrality of Visual Product Aesthetics questionnaire (CVPA; Bloch et al., 2003; German version Thielsch, 2008). Afterwards, participants were asked to assess the aesthetics of a given webpage.

The experiment started with 10 practice trials. There was a short break after the practice trials in which ambiguities on the participant's side could be solved. A fixation cross was presented to 250 ms in screen center, then light-grey screen was displayed for 1250 ms, followed by displaying the webpage-screenshot for 2500 ms. Again, a light grey screen was displayed for 500 ms when three exclamation marks appeared to signal the participants to give their aesthetic judgment (aesthetic or unaesthetic) by pressing one of two keys. There were short breaks of 10 – 15 seconds duration every 90 – 120 seconds and a longer break of 90 – 135 seconds every 7.5 – 9.0 minutes. The experiment lasted about 25 minutes. Stimuli, 750 in total, were presented in random order determined for each participant.

## Data analysis

The EEG-data were rereferenced to linked mastoids. An offline bandpass-filter (half-power: .1 – 25 Hz) was applied. An EEG larger than  $\pm 75 \mu V$  was considered an artifact. Artifact free epochs of 1200 ms length with a baseline of 200 ms were defined. A trial was considered congruent if there was a congruency of the participant's aesthetic judgment and our conditioning in aesthetic and unaesthetic webpages. This resulted in 705 aesthetic (experts: 292, laypersons: 413) and 1136 unaesthetic congruent trials (experts: 541, laypersons: 595). The

**Comment [A30]:** Did this include a fixation cross as well (as the previous screen did) what I am asking is – were participants encouraged not to move their eyes- before the crucial display appeared. Were participants already centrally fixated before the website display appeared on screen?

**Comment [A31]:** A similar question- because the stimuli were shown for a relatively long period (in EEG terms) it might be good to mention that you encouraged participants to maintain central fixation – they knew they would be making an evaluation and you would have been encouraging them to use covert attention rather than overt visual search strategy by moving eyes during evaluation? Just need some mention of how you dealt with this potential issue

**Comment [A32]:** This really does need clarification- at no point do the authors specify how there are 750 trials/stimuli in the experiment- we are given the information that there are 75 stimuli in each of the two sets (150 stimuli in total) I presume that these are repeated across blocks? so that there are a total of 750 presentations- if this is the case you need to report whether each set is randomly presented within a block (and therefore repeated- need to report how many times stimuli are repeated) For example... participants see 150 random presentations of all the stimuli followed by another 150 random presentations and so on Or is it as you state that "...750 in total, were presented in random order determined for each participant." This would mean that the same website image could be repeated a number of times in quick succession – this is very important and needs to be made clear If the two sets were repeated across the experiment did repetition interact with aesthetic/unaesthetic – basically if participants didn't like a website on view 1 did they not like it still after it had been repeated a number of times? I think this is a point that needs some clarification and doesn't need to be complicated

**Comment [A33]:** The term 'conditioning' in this context does not make sense I presume you mean that Only those trials where participants classification of aesthetic or unaesthetic agreed with the classification from the pilot study were included in the grand average event related potentials for each of the two conditions?? You need to make this clearer This may of course lead to differences in S/N ratio but Steven Luck suggests that ERP analyses is relatively robust do difference in s/n so not necessarily a problem

remaining trials were excluded due to incongruency or artifacts. We calculated mean voltage per participant and condition in the following time-windows 80 ms – 105 ms, 110 ms – 130 ms, 150 ms – 370 ms, 370 ms – 600 ms and 600 ms – 800 ms after stimulus onset. Electrodes were grouped into lateral-central position (left: F3, C3, P3, O1; central: Fz, Cz, Pz, Oz; right: F4, C4, P4, O2), and anterior-posterior position (frontal: F3, Fz, F4; central: C3, Cz, C4; parietal: P3, Pz, P4; occipital: O1, Oz, O2). Mean voltages per participants and time-windows were subjected to mixed ANOVAS (repeated measurement factors: Webpage Aesthetics: aesthetic-unaesthetic; Electrode Group: lateral-central; anterior-posterior; between-group factor: Group: expert vs. layperson).

**Comment [A34]:** This really needs clarification and is related to the issue of number of trials- repeats etc., Does this represent the number of trials in the grand average ERPs??? Is so are you saying that the signal to noise ratio differs between conditions I think this section needs clarification

**Comment [A35]:** Was there theoretical reason for all these time windows? Check this out

**Comment [A36]:**

**Comment [A37]:** I presume these are two different factors? If as below please correct  
Factors Aesthetics 2 (aesthetic/unaesthetic)  
LateralCentral 3(left/mid/right) PostAnt 2(anterior/posterior).....

## Results

### Behavioral Data

Experts and laypersons evaluated unaesthetic webpages similarly but differed in their evaluations concerning aesthetic webpages (see Table 1). A generalized linear mixed logistic regression model was calculated next (Barr, Levy, Scheepers & Tily, 2013; Jaeger, 2008).

Table 1: Frequency of aesthetic evaluations as a function of conditions

group	Evaluation			
	aesthetic	unaesthetic	aesthetic	unaesthetic
	unaesthetic webpage		aesthetic webpage	
expert	94	656	355	395
layperson	97	653	458	292

**Comment [A38]:** As participants made a behavioral response during the recording of EEG it would be useful to see the mean RTs for the two groups under the two different stimulus types- importantly because the authors report ERPs during the P3 this component is usually correlated with response time- faster response times elicit shorter latency P3 (and associated amplitude differences that result from this)

**Comment [A39]:** remove

The factors Group (expert vs. layperson) and Webpage Aesthetics (aesthetic vs. unaesthetic) served as fixed effects. Trial-number, participant and webpage served as random effects. Random slope and random intercepts for participants and webpage were realized. Dependent variable was the aesthetic evaluation by the participant. The stepwise inclusion of the predictor Webpage Aesthetics and the interaction of Group and Webpage Aesthetics resulted in a significant increase of fit ( $\chi^2(7) = 97.58, p < .001$ ;  $\chi^2(8) = 8.09, p = .005$ , respectively, see Table 2 for a summary). The Odds ratio for the predictor Webpage Aesthetics changes by a factor of 13.00 as the webpage aesthetics changes from aesthetic to unaesthetic. The significant interaction shows that the Odds Ratio for an aesthetic evaluation changes by a factor of 1.88 depending on group membership but only if an aesthetic webpage is being evaluated.

Table 2: Summary regression analysis

effect	$\beta$	SE( $\beta$ )	Odds Ratio	<i>p</i> -value
Intercept	-2.30	.24	0.10	< .001
Group	-.03	.23	.97	
Webpage Aesthetics	2.56	.28	12.94	< .001
Group x Webpage Aesthetics	.63	.22	1.88	< .01

In sum, experts evaluated aesthetic webpages more often as unaesthetic compared to laypersons.

### EEG-Data

The continuous EEG-signal was split up in five different time windows: 80 ms – 105 ms (N1), 110 ms – 130 ms (P1), 150 ms – 370 ms (N2), 370 ms – 600 ms and 600 ms – 800 ms after stimulus onset. We used mean amplitude of each time window as dependent variable in separate

mixed ANOVAs with averaged Electrode Position (left, central, right and frontal, central, parietal, occipital see above Data Analysis), Webpage (aesthetic, unaesthetic) as repeated measurements factors and Group (expert, layperson) as between factor. Greenhouse-Geisser corrected degrees of freedom are reported in case of a sphericity assumption violation. We only report (near) significant results of experimentally manipulated factors to allow for an easier overview of the results. Figure 1 and Figure 2 show the EEG-signal averaged in the regions of interest (Figure 1: anterior, central, parietal, occipital; Figure 2: left, central, right).

**Comment [A40]:** see previous comment re factors

**Comment [A41]:** Grand average ERPs. Is the resolution of these figures sufficient for publication? It could be just the copies I have are not that great?

#### Time window 80 ms – 105 ms after stimulus onset

There is an N1 with a maximum at 90 ms. The main effect of Group fails significance levels ( $F(1, 18) = 4.28, p = .053, \eta_p^2 = .19$ ). Neither the main effect Webpage nor the interaction of Group and Webpage are significant ( $F(1, 18) = 1.52, p = .233, \eta_p^2 = .08$ ;  $F(1, 18) = 1.07, p = .314, \eta_p^2 = .06$ ). There are no further significant interactions with Electrode Position.

**Comment [A42]:** I'm not sure how you can call this an N1 – it might be an N1 at one of the electrode clusters you analyse but not all of them – please see the next comment

**Comment [A43]:** Maximum amplitude? Maximum difference? Be clear. Also how can there be a max at 90ms as the peak will differ across scalp locations?

#### Time window 110 ms – 130 ms after stimulus onset

In this time a P1 with a maximum at 120 ms appeared. Aesthetically evaluated webpages resulted in a more negative going activation than unaesthetically evaluated webpages ( $F(1, 18) = 4.49, p = .048, \eta_p^2 = .20$ ). Laypersons have a more positive ERP than experts in this time window ( $F(1, 18) = 4.98, p = .039, \eta_p^2 = .22$ ). There are no significant interactions.

**Comment [A44]:** I think you will need to provide some rationale for the analyses of the time windows you focus on? What is this analyses trying to capture for example – If you are looking at early sensory/perceptual processing and are interested in the mean amplitude of the P1 then why are you analyzing erps across the whole head during this time window? You could have saved a lot of analyses by focusing on those component of interest (that have a theoretical fit with the model you are testing the neural plausibility)

**Comment [A45]:** P1?

#### Time window 150 – 370 ms after stimulus onset

There is distinct negative going ERP-wave with its maximum at 240 ms after stimulus onset. The main effects of Group and Webpage are both significant while the interaction of Group and Webpage is not significant ( $F(1, 18) = 4.82, p = .042, \eta_p^2 = .21$ ;  $F(1, 18) = 32.43, p = .001, \eta_p^2 = .64$ ; ( $F(1, 18) = 2.33, p = .145, \eta_p^2 = .11$ ; respectively). Experts and aesthetic websites

**Comment [A46]:** Can you provide (all p's > \* or F's < I realize that you have not reported non-significant results for brevity but this is very unusual – could you at least show how close your effects were to significance

**Comment [A47]:** No this is a posterior positivity and anterior negativity in the ERP waveforms -

**Comment [A48]:** Why do you report the statistics for the non-significant interaction between group and webpage in this time window but not in the previous time window which showed a similar effect?

evoke stronger negative ERPs. There are significant interactions of Webpage, lateral-central and anterior-posterior Electrode Positions ( $F(2, 36) = 3.31, p = .048, \eta_p^2 = .15$ ;  $F(1.44, 25.99) = 17.22, p < .001, \eta_p^2 = .49$ ). All other interactions are not significant.

Follow up analysis with Webpage and Electrode Position serving as factors shows significant effects of Webpage for each lateral-central Electrode Position (left:  $F(1, 18) = 29.41, p < .001, \eta_p^2 = .62$ ; right:  $F(1, 18) = 32.33, p < .001, \eta_p^2 = .64$ ; central:  $F(1, 18) = 31.04, p < .001, \eta_p^2 = .63$ ). Lateral-central Electrode Position on Webpage were also significant (aesthetic:  $F(2, 36) = 22.45, p < .001, \eta_p^2 = .56$ ; unaesthetic:  $F(2, 36) = 11.69, p < .001, \eta_p^2 = .39$ ). Concerning the other significant interaction between Webpage and anterior-posterior Electrode Position, follow up analyses reveals significant effects of Webpage on frontal, central and parietal but not on occipital electrodes (frontal:  $F(1, 18) = 39.01, p < .001, \eta_p^2 = .68$ ; central:  $F(1, 18) = 44.78, p < .001, \eta_p^2 = .71$ ; parietal:  $F(1, 18) = 26.80, p < .001, \eta_p^2 = .60$ ; occipital:  $F(1, 18) = .66, p = .426, \eta_p^2 = .04$ ). There was also a significant effect of Webpage at anterior-posterior Electrode Position (aesthetic:  $F(1.29, 23.20) = 33.72, p < .001, \eta_p^2 = .65$ ; unaesthetic:  $F(1.38, 24.86) = 25.19, p < .001, \eta_p^2 = .58$ ).

#### **Time window 370 – 600 ms after stimulus onset**

In this time window, the EEG-amplitudes are less pronounced than in the previous time windows and are getting more positive. As before, aesthetic webpages exhibit a more negative EEG as unaesthetic webpages ( $F(1, 18) = 6.60, p = .019, \eta_p^2 = .27$ ). Experts show a more negative EEG than laypersons ( $F(1, 18) = 6.22, p = .023, \eta_p^2 = .26$ ). The interaction of Webpage and anterior-posterior Electrode Position is also significant ( $F(1.51, 27.1) = 9.98, p = .001, \eta_p^2 = .36$ ). The remaining interactions are not significant.



Follow up analysis shows that aesthetic webpages differ from unaesthetic webpages at frontal ( $F(1, 18) = 6.82, p = .018, \eta_p^2 = .28$ ), central ( $F(1, 18) = 19.60, p < .001, \eta_p^2 = .52$ ) and parietal position ( $F(1, 18) = 4.85, p = .041, \eta_p^2 = .21$ ), but not at occipital electrode positions ( $F < 1$ ). Moreover, there is a significant effect for anterior-posterior Electrode Position for aesthetic as well as unaesthetic Webpages ( $F(1.41, 25.39) = 6.74, p = .009, \eta_p^2 = .27$ ;  $F(1.64, 29.57) = 5.56, p = .013, \eta_p^2 = .24$ ).

#### Time window 600 – 800 ms after stimulus onset

In this time window, the EEG comes from a negative voltage range into a positive one at frontal and central electrodes while the EEG at occipital electrodes shows the opposite patterns. The EEG at parietal electrodes fluctuates around zero.

There are no main effects, but a significant interaction of Webpage and anterior-posterior Electrode Position ( $F(1.87, 33.58) = 4.00, p = .030, \eta_p^2 = .18$ ). Follow up analysis on this interaction reveals a significant effect of anterior-posterior Electrode position for aesthetic as well as unaesthetic Webpages ( $F(1.47, 26.44) = 4.61, p = .028, \eta_p^2 = .20$ ;  $F(1.67, 30.02) = 7.69, p = .003, \eta_p^2 = .30$ ).

#### Discussion

We asked design-experts and laypersons to evaluate the aesthetic properties of static webpages that varied in aesthetic attractiveness. Behavioral, is the webpage attractive or unattractive, and electrophysiological responses were registered. We will discuss the behavioral data next, followed by discussing the electrophysiological data. The remaining section relates the results to the models presented above.

**Comment [A49]:** Was reaction time recorded? Along with the 1 -1 response???

**Comment [A50]:** recorded

**Comment [A51]:** sentence

Participants' responses indicated that they were more critical than anticipated. Although we presented an even number of aesthetic and unaesthetic webpages, unaesthetic judgements prevailed. Laypersons and experts evaluate unaesthetic webpages similarly. In case of aesthetic webpages, experts seemed to be more critical than laypersons given that they evaluated more webpages as unaesthetic than laypersons did. Nonetheless, the major factor driving the evaluation is the aesthetic quality of the webpage itself, not the person evaluating the webpage as indicated by the 6-7 times larger odds ratio for webpage aesthetic than the odds-ratio for group membership. Given this, expertise is less important than aesthetics in evaluating a webpage.

There were only numerical differences in the N1-window with experts exhibiting a non-significant more negative N1. This indication of a difference between experts and laypersons becomes evident in the following time windows. Aesthetically evaluated webpages result in a stronger N1 than unaesthetically evaluated ones. At the same time, experts exhibit a more negative EGG than laypersons. Interestingly, there were no interactions between group and webpage in the investigated time range. Thus, the difference in processing seems to start early carrying on to later processing.

In terms of scalp distribution, the effect seems rather stable. There is an interaction between electrode position and aesthetic evaluation starting at around 150 ms and lasting to 600 ms after stimulus onset. An aesthetic evaluation results in a more negative evaluation at frontal, central and parietal electrodes than at occipital electrodes. This pattern reverses independently of aesthetic evaluation in a time window lasting from 600 – 800 ms. At the same time, the long-lasting temporal and large spatial distribution of the effects indicate that various, interconnected cortical areas are involved in aesthetic processing (Kawabata & Zeki, 2004). Distinct processing stages could not be identified here.

**Comment [A52]:** please see the notes in the results section  
I'm not sure what N1 means in this context as N1 refers to an ERP component that will have differing functional characteristics depending on its spatial location- during this time window you included posterior P1) amplitudes so this doesn't make sense- you would be better to simply refer to the time window ie 80-105 ms across the whole head  
I hope this is clear

**Comment [A53]:** Non significant what? Difference in trials comparing?

**Comment [A54]:** EGG?

**Comment [A55]:** What does this mean  
Early sensory differences- it is difficult to interpret your ERP components because you have not related them to previously reported components ie frontal N1 posterior P1 doing so will make your effects more straightforward to interpret

**Comment [A56]:** What does this mean?  
Do you mean they are long latency effects rather than transient short lived differences (ie significant difference are found will a similar direction across a number of time windows)

**Comment [A57]:** this does not make sense-  
if you have clear predictions about specific components (that are functionally related to your model) then you can talk about the processing stages  
for example the EDAN and LPP have particular functional attributes that you can relate to your design (and make very specific predictions in relation to these)

**Experts versus laypersons**

Modern theories of aesthetic processing assume an interaction of perceiver and object such that experts should process objects differently than laypersons. Our data partly support this assumption. Experts were more critical than laypersons as indicated by the behavioral data. However, a similar interaction of expertise and aesthetic evaluation, for instance greater group differences for aesthetic webpages than for unaesthetic webpages, is absent in the ERPs. This pattern would have been predicted by recent theories of aesthetic processing (Chatterjee, 2004; Leder et al., 2004; Reber et al., 2004). The main effects of Group and Webpage suggest that the neuronal generators underlying the ERPs are basically the same but experts process the stimuli in a more thorough manner.

Differences between experts and laypersons have been observed in various behavioral studies (e.g., Hekkert et al., 1994; Hekkert & van Wieringen, 1996a; Hekkert & van Wieringen, 1996b; Phiko et al., 2011; Winston & Cupchik, 1992). For instance, it has been suggested that experts have enhanced associative knowledge that is easily accessible (Cheung & Bar, 2012; Harel et al., 2010; Long et al., 2011, Tanaka & Taylor, 1991). Such enhanced memory representations might be activated when processing the presented webpages. Furthermore, motivational and attentional differences might contribute to the observed difference. Harel et al. (2010) showed that expertise influences neuronal activation mainly when the expertise is task-relevant. Therefore, it is not a bottom-up, stimulus-driven processing mechanism that differs between experts and laypersons but rather a top-down (i.e. task demands) modulated processing that results in intensified processing. If this explanation holds, experts and laypersons should not differ when the expertise is not task-relevant as in the current study.

### **Aesthetics of webpages**

Attempts to obtain an objective measure of the aesthetic value of an object can be found in Altaboli and Lin (2011), Ngo, Samsudin, and Abdullah (2000) or Seckler, Opwis, and Tuch, (2015). We used dimensions such as simplicity, versatility, colorfulness, and artistry to divide the webpages in aesthetic and unaesthetic ones (Moshagen & Thielsch, 2010). Nonetheless, we are not able to determine which of these dimensions (or which combination of dimensions) brought about the observed behavioral and electrophysiological differences between aesthetic and unaesthetic webpages.

Aesthetic and unaesthetic webpages elicit different EEG-amplitude in an early time-window of 110 – 130 ms. ERPs in this time window presumably reflect processing of stimulus properties such as contrast or brightness (e.g. Luck, 2005). Often such early visual processes are expressed most over occipital electrodes, which is not what we observe here. We observe a frontal to parietal distribution sparing occipital electrodes. This distributional pattern suggests that not visual properties such as contrast or brightness brought about these differences. However, aesthetic and unaesthetic webpages differed in complexity measured in bytes with unaesthetic webpages being more complex than aesthetic ones. High visual complexity usually results in a more negative evaluation than medium to less complexity (Tuch et al., 2012a). Therefore, this early differentiation might reflect visual complexity. But keep in mind, that our complexity measure might be rather crude and probably does not reflect functional complexity. Other properties could bring about the observed difference.

For instance, it might be that (aesthetic) webpages exhibit fractal-like image properties as do graphic art or natural scenes (Redies, Hasenstein & Denzler, 2007). Redies et al. (2007) link such fractal-like image properties to the aesthetic perception. A theory of aesthetic processing

must take into account human sensory processing. Consequently, web-designer and artists exploit such image-properties because the human visual system has evolved that way. It remains to be determined whether webpages exhibit fractal-like properties and whether differentiate between aesthetic and unaesthetic webpages. Keep in mind that fractal-like image properties are probably one of many properties contributing to an aesthetic evaluation. Whether they are necessary or sufficient for an aesthetic evaluation needs to be determined (Redies et al., 2007).

The early differentiation we observed here is in contrast to theories put forward by Chatterjee (2004) or Leder et al. (2004) who do not assume an influence of aesthetic properties on early processing. More in line with these assumptions are the results by Höfel and Jacobsen (2007). They observed a differentiation of ERPs to aesthetic and unaesthetic picture not earlier than 300 ms after stimulus onset (see also Jacobsen & Höfel, 2003 using the same stimuli but a different task for a similar time pattern). The stimuli, Höfel and Jacobsen (2007) employed, were black-white symmetric and asymmetric patterns instead of colored webpages as we used. Such stimuli apparently elicit similar ERP-patterns but they might miss properties that the webpages employed here had, for instance, different degree of complexity, color and so on.

Some studies in this area (Lindgaard et al., 2006, 2011; Tractinsky, 2006; etc.) have presented stimuli for 50 ms (Tuch et al., 2012a even for 17 ms) and evaluations of these shortly presented stimuli were quite stable – but it is not to be supposed that the cognitive processing of these stimuli only takes 50 ms. Based on our data we assume that evaluation is a rather long lasting process which can be initiated even with brief presentation durations.

We argued above that experts have enhanced, widely distributed representations that are easy to access. Thus, attention processes operating in a form of a feed-forward sweep might have influenced aesthetic evaluation processes already early on (Chatterjee, 2004). However, this

would imply an interaction of expertise and webpage aesthetics that was absent here. The aesthetic evaluation is probably based on a variety of stimulus properties that are processed in a bottom-up manner first (Douneva, Jaron & Thielsch, in press; Thielsch & Hirschfeld, 2012) before top-down processes kick in. It is rather unlikely, that bottom-up processing affects stimulus processing over the whole analysis period. Rather, evaluative impression formation and evaluative categorization take place in this time period (Cela-Conde et al., 2004; Jacobsen & Höfel, 2003). However, the long lasting difference between aesthetic and unaesthetic webpages prohibits relating cognitive processes to time periods in a fine-grained manner. The spatial distribution of the ERP-effect is of little help. Spatial and temporal distributions of ERP-effects allow only relatively gross classification. Evaluative processes might bring about the interaction of aesthetic and anterior-posterior activation in the time-window of 150 ms – 600 ms. The occipital activation in the earlier time-window might reflect mainly processing of aesthetic stimulus properties.

Aesthetic webpages elicited more negative ERP than unaesthetic ones. There is a more negative going ERP at central to lateral electrodes in a time-window of 150 ms to 370 ms after stimulus onset for aesthetic compared to unaesthetic webpages. Cela-Conde et al. (2004) but also Jacobsen and Höfel (2003) observed much more temporally and spatially distinct differences between aesthetic and unaesthetic stimuli. Jacobsen and Höfel observed a more negative going ERP when participants viewed unaesthetic stimuli. However, they compared “beautiful” decisions to “symmetric”-decision. The different temporal and spatial distribution might be due to the employed task. In addition, while we used linked mastoids as reference electrode, Jacobsen and Höfel used the nose tip as reference electrode. That prevents rereferencing our data to their setup. Thus, the differences in spatial distribution might simply be due to reference differences.

Cela-Conde et al. (2004) observed effects starting at around 400 – 900 ms after stimulus onset while participants observed aesthetic (rated beautiful) stimuli. The late onset might be due to using artistic and non-artistic stimuli instead of “everyday” aesthetic and unaesthetic stimuli. Furthermore, the spatial differences probably result from the fact that Cela-Conde et al. used MEG, a reference-free measure, and equivalent dipoles in the source space to determine the spatial distribution. The “relatively” small number of electrodes that we used prohibits source-location. In sum, task, stimuli and recording technique might contribute to the observed differences.

## **Conclusions**

The relevant aesthetic theories (Chatterjee, 2004; Leder et al., 2004; Reber et al., 2004) predict an interaction between recipient characteristics and stimulus properties. Leder et al. assume that an aesthetic form should facilitate perceptual and cognitive processing given expertise (see also Reber et al., 2004). More expertise should result in less cognitive effort, hence in less neuronal activation. We could not find such interaction; rather we observed only main effects of recipient characteristics and of stimulus properties. Also not anticipated, experts showed more activation than laypersons. One might assume, that the observed activation reflects the broader and better-connected associative network that experts supposedly develop (Cheung & Bar, 2012; Harel, et al., 2010). Nonetheless, differences between aesthetically and unaesthetically judged webpages emerge much earlier than anticipated.

### References

- Altaboli, A. & Lin, Y. (2011). Objective and Subjective Measures of Visual Aesthetics of Website Interface Design: The Two Sides of the Coin. In J.A. Jacko (Ed.): *Human-Computer Interaction* (pp. 35–44), Heidelberg: Springer.
- Arnheim, R. (1974). *Art and visual Perception: A psychology of the creative eye*. Berkeley: University of California Press.
- Barr, D. J., Levy, R., Scheepers, C. & Tily, H. J. (2013). Random effect structures for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language*, 68, 255 – 278. doi: 10.1016/j.jml.2012.11.001
- Bentin, Sh, Allison, T., Puce, A. Perez, E. & McCarthy, G. (1996). Electrophysiological Studies of Face Perception in Humans. *Journal of Cognitive Neuroscience*, 8, 551-565. doi:10.1162/jocn.1996.8.6.551
- Berlyne, D. E. (1971). *Aesthetics and Psychobiology*. New York: Appleton-Century-Crofts.
- Birkhoff, G. D. (1933). *Aesthetic measure*. Cambridge: Harvard University Press
- Bloch, P. H., Brunel, F. F. & Arnold, T. J. (2003). Individual differences in the centrality of visual product aesthetics: Concept and measurement. *Journal of Consumer Research*, 29, 551- 565.
- Carretié, L., Mercado, F., Tapia, M. & Hinojosa, J. A. (2001). Emotion, attention and the 'negativity bias', studied through event-related potentials. *International Journal of Psychophysiology*, 41(1), 75-85. doi:10.1016/S0167-8760(00)00195-1



- Cela-Conde, C. J., Marty, G., Maestú, F., Ortiz, T., Munar, E., Fernández, A., ... Quesney, F. (2004). Activation of the prefrontal cortex in the human visual aesthetic perception. *Proceedings of the National Academy of Sciences of the United States of America*, 101(16), 6321–6325. <http://doi.org/10.1073/pnas.0401427101>
- Chatterjee, A. (2004). Prospects for a Cognitive Neuroscience of Visual Aesthetics. *Bulletin of Psychology and the Arts*, 4(2), 55-60. doi: 10.1037/e514602010-003
- Cheung, O. S. & Bar, M. (2012). Visual prediction and perceptual expertise. *International Journal of Psychophysiology*, 83, 156-163. doi: 10.1016/j.ijpsycho.2011.11.002
- Chevalier, A. & Ivory, M. Y. (2003). Web site designs: Influences of designer's expertise and design constraints. *International Journal of Human - Computer Studies*, 58, 57–87. doi: 10.1016/S1071-5819(02)00126-X
- de Tommaso, M., Pecoraro, C., Sardaro, M., Serpino, C., Lancioni, G. & Livrea, P. (2008). Influence of aesthetic perception on visual event-related potentials. *Consciousness and Cognition*, 17, 933–945. doi:10.1016/j.concog.2007.09.003
- Douneva, M., Jaron, R. & Thielsch, M.T. (in press). Effects of different website designs on first impressions, aesthetic judgments, and memory performance after short presentation. *Interacting with Computers*. doi: 10.1093/iwc/iwv033
- Eysenck, H. (1941). The empirical determination of an aesthetic formula. *Psychological Review*, 48, 83–92.
- Fechner G. T. (1876). *Vorschule der Aesthetik [Experimental Aesthetics; "Pre-school" of aesthetics]*. Leipzig: Breitkopf & Härtel

- Gauthier, I., Skudlarski, P., Gore, J., Anderson, A. W. (2000). Expertise for cars and birds recruits brain areas involved in face recognition. *Nature Neuroscience*, 3, 191-197. doi:10.1038/72140
- Gauthier, I. & Tarr, M. J. (1997). Becoming a “Greeble” Expert: Exploring Mechanisms for Face Recognition. *Vision Research*, 37, 1673-1682. doi:10.1016/S0042-6989(96)00286-6
- Gombrich, E. H. (1995). *The Story of art*. London: Phaidon.
- Handy, T. C., Smilek, D., Geiger, L., Liu, C. & Schooler, J. W. (2010). ERP evidence for rapid hedonic evaluation of logos. *Journal of Cognitive Neuroscience*, 22(1), 124-138.
- Harel, A., Gilaie-Dotan, S., Malach, R. & Bentin, S. (2010). Top-down engagement modulates the neural expressions of visual expertise. *Cerebral Cortex*, 20(10), 2304-2318. doi: 10.1093/cercor/bhp316
- Hasse, C. & Weber, R. (2012). Eye movements on facades: The subjective perception of balance in architecture and its link to aesthetic judgment. *Empirical Studies of The Arts*, 30(1), 7-22. doi:10.2190/EM.30.1.c
- Hekkert, P. & van Wieringen, P. C. W. (1990). Complexity and prototypicality as determinants of the appraisal of cubist paintings. *British Journal of Psychology*, 81, 483-495. doi: 10.1111/j.2044-8295.1990.tb02374.x
- Hekkert, P., Snelders, D. & Wieringen, P. C. W. (2003). ‘Most advanced, yet acceptable’: Typicality and novelty as joint predictors of aesthetic preference in industrial design. *British Journal of Psychology*, 94, 111-124. doi: 10.1348/000712603762842147

- Hekkert, P., & van Wieringen, P. C. (1996a). The impact of level of expertise on the evaluation of original and altered versions of post-impressionistic paintings. *Acta Psychologica*, 94, 117-131. doi:10.1016/0001-6918(95)00055-0
- Hekkert, P., & Van Wieringen, P. C. (1996b). Beauty in the eye of expert and nonexpert beholders: A study in the appraisal of art. *The American Journal of Psychology*, 109, 389-407. doi: 10.2307/1423013
- Höfel, L. & Jacobsen, T. (2007a). Electrophysiological indices of processing aesthetics: Spontaneous or intentional processes? *International Journal of Psychophysiology*, 65(1), 20-31. doi: 10.1016/j.ijpsycho.2007.02.007
- Höfel, L. & Jacobsen, T. (2007b). Electrophysiological indices of processing symmetry and aesthetics. *Journal of Psychophysiology*, 21(1), 9-21. doi: 10.1027/0269-8803.21.1.9
- Hoyer, W. D. & Stokburger-Sauer, N. E. (2012). The role of aesthetic taste in consumer behavior. *Journal of The Academy of Marketing Science*, 40(1), 167-180. doi:10.1007/s11747-011-0269-y
- Huang, Y. X., & Luo, Y. J. (2006). Temporal course of emotional negativity bias: an ERP study. *Neuroscience letters*, 398, 91-96. doi:10.1016/j.neulet.2005.12.074
- ISO (2006). ISO 9241: Ergonomics of Human-System Interaction – Part 151: Guidance on World Wide Web Interfaces. Geneva: International Organisation for Standardisation.
- Jacobsen, T. & Höfel, L. (2003). Descriptive and evaluative judgment processes: Behavioral and electrophysiological indices of processing symmetry and aesthetics. *Cognitive, Affective & Behavioral Neuroscience*, 3(4), 289-299. doi: 10.3758/CABN.3.4.289

- Jacobsen, T. (2006). Bridging the Arts and Sciences: A Framework for the Psychology of Aesthetics. *Leonardo*, 39, 155-162.
- Jacobsen, T. (2010). Beauty and the brain: culture, history and individual differences in aesthetic appreciation. *Journal of Anatomy*, 216(2), 184-191
- Jaeger, T. F. (2008). Categorical data analysis: Away from ANOVAs (transformations or not) and towards logit mixed models. *Journal of Memory and Language*, 59, 434 – 446. doi: 10.1016/j.jml.2007.11.007
- Kawabata, H. & Zeki, S. (2004). Neural correlates of beauty. *Journal of Neurophysiology*, 91, 1699-1705. doi: 10.1152/jn.00696.2003
- Kirk, U., Skov, M., Christensen, M. S., & Nygaard, N. (2009). Brain correlates of aesthetic expertise: a parametric fMRI study. *Brain and Cognition*, 69, 306-315. doi:10.1016/j.bandc.2008.08.004
- Kowatari, Y., Lee, S. H., Yamamura, H., Nagamori, Y., Levy, P., Yamane, S., & Yamamoto, M. (2009). Neural networks involved in artistic creativity. *Human Brain Mapping*, 30, 1678-1690. doi: 10.1002/hbm.20633
- Kozbelt, A., Dexter, S., Dolese, M. & Seidel, A. (2012). The aesthetics of software code: A quantitative exploration. *Psychology of Aesthetics, Creativity, and The Arts*, 6(1), 57-65. doi:10.1037/a0025426 doi: 10.1037/a0025426
- Lavie, T. & Tractinsky, N. (2004). Assessing Dimensions of Perceived Visual Aesthetics of Web Sites. *International Journal of Human - Computer Studies*, 60, 269–298. doi: 10.1016/j.ijhcs.2003.09.002

- Leder, H. & Nadal, M. (2014). Ten years of a model of aesthetic appreciation and aesthetic judgments : The aesthetic episode - Developments and challenges in empirical aesthetics. *British Journal of Psychology*, 105(4), 443–464. doi:10.1111/bjop.12084
- Leder, H., Belke, B., Oeberst, A. & Augustin, D. (2004). A model of aesthetic appreciation and aesthetic judgements. *British Journal of Psychology*, 95, 489–508. doi: 10.1348/0007126042369811
- Lee, S. & Koubek, R. J. (2012). Users' perceptions of usability and aesthetics as criteria of pre- and post-use preferences. *European Journal of Industrial Engineering*, 6(1), 87. doi: 10.1504/EJIE.2012.044812
- Long, Z., Peng, D., Chen, K., Jin, Z., & Yao, L. (2011). Neural substrates in color processing: A comparison between painting majors and non-majors. *Neuroscience Letters*, 487, 191-195. doi:10.1016/j.neulet.2010.10.020
- 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*, 18(1), 1–30. doi: 10.1145/1959022.1959023
- 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. doi: 10.1080/01449290500330448
- Luck, S. J. (2005). *An Introduction to the Event-related Potential Technique*. Cambridge, MA. MIT Press.
- Marr, D. (1982). *Vision: A computational investigation into the human representation and processing of visual information*. New York: WH Freeman and Company.

- Moshagen, M. & Thielsch, M. T. (2010). Facets of visual aesthetics. *International Journal of Human-Computer Studies*, 68 (10), 689-709. doi: 10.1016/j.ijhcs.2010.05.006
- Moshagen, M. & Thielsch, M. T. (2013). A short version of the visual aesthetics of websites inventory. *Behaviour & Information Technology*, 32 (12), 1305-1311. doi: 10.1080/0144929X.2012.694910
- Müller, M., Höfel, L., Brattico, E. & Jacobsen, T. (2009), Electrophysiological Correlates of Aesthetic Music Processing. *Annals of the New York Academy of Sciences*, 1169, 355–358. doi: 10.1111/j.1749-6632.2009.04846.x
- Müller, M., Höfel, L., Brattico, E. & Jacobsen, T. (2010). Aesthetic judgments of music in experts and laypersons—An ERP study. *International Journal of Psychophysiology*, 76(1), 40-51. doi:10.1016/j.ijpsycho.2010.02.002
- Ngo, D. C. L., Samsudin, A. & Abdullah, R. (2000). Aesthetic measures for assessing graphic screens. *Journal of Information Science and Engineering*, 16(1), 97–116.
- Pang, C. Y., Nadal, M., Müller-Paul, J. S., Rosenberg, R., & Klein, C. (2013). Electrophysiological correlates of looking at paintings and its association with art expertise. *Biological Psychology*, 93, 246-254. doi:10.1016/j.biopsycho.2012.10.013
- Park, S., Choi, D. & Kim, J. (2004). Critical factors for the aesthetic fidelity of web pages: empirical studies with professional web designers and users. *Interacting with Computers*, 16, 351–376. doi:10.1016/j.intcom.2003.07.001
- Pihko, E., Virtanen, A., Saarinen, V. M., Pannasch, S., Hirvenkari, L., Tossavainen, T., ... & Hari, R. (2011). Experiencing art: the influence of expertise and painting abstraction level. *Frontiers in Human Neuroscience*, 5, 1-10. doi: 10.3389/fnhum.2011.00094

- 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. doi:10.1207/s15327957pspr0804\_3
- Redies, C., Hasenstein, J. & Denzler, J. (2007). Fractal-like image statistics in visual art: similarity to natural scenes. *Spatial Vision*, 21(1), 137-148. doi: 10.1163/156856808782713825
- Scott, L. S., Tanaka, J. W., Sheinberg, D. L., & Curran, T. (2006). A reevaluation of the electrophysiological correlates of expert object processing. *Journal of Cognitive Neuroscience*, 18, 1453-1465. doi: 10.1162/jocn.2006.18.9.1453
- Seckler, M., Opwis, K. & Tuch, A. N. (2015). Linking objective design factors with subjective aesthetics: An experimental study on how structure and color of websites affect the facets of users' visual aesthetic perception. *Computers in Human Behavior*, 49, 375-389. doi:10.1016/j.chb.2015.02.056
- Silvia, P. J. & Berg, C. (2011). Finding movies interesting: How appraisals and expertise influence the aesthetic experience of film. *Empirical Studies of the Arts*, 29, 73-88. doi:10.2190/EM.29.1.e
- Silvia, P. J. (2013). Interested experts, confused novices: Art expertise and the knowledge emotions. *Empirical Studies of the Arts*, 31, 107-115. doi: 10.2190/EM.31.1.f
- Solso, R. L. (2003). *The psychology of art and the evaluation of the conscious brain*. Cambridge: MIT Press.
- Tanaka, J. W., Curran, T. (2001). A Neural Basis for Expert Object Recognition. *Psychological Science*, 12, 43-47. doi: 10.1111/1467-9280.00308

- Tanaka, J. W., & Taylor, M. (1991). Object categories and expertise: Is the basic level in the eye of the beholder?. *Cognitive Psychology*, 23, 457-482. doi:10.1016/0010-0285(91)90016-H
- Thielsch, M. (2008). Ästhetik von Websites: Wahrnehmung von Ästhetik und deren Beziehung zu Inhalt, Usability und Persönlichkeitsmerkmalen (*Aesthetics of websites: Perception of aesthetics and its relation to content, usability and traits*). Münster: MV Wissenschaft.
- Thielsch, M. T. & Hirschfeld, G. (2012). Spatial frequencies in aesthetic website evaluations – explaining how ultra-rapid evaluations are formed. *Ergonomics*, 55 (7), 731-742. doi: 10.1080/00140139.2012.665496
- Thielsch, M. T., Blotenberg, I. & Jaron, R. (2014). User evaluation of websites: From first impression to recommendation. *Interacting with Computers*, 26, 89-102. doi: 10.1093/iwc/iwt033
- 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. doi: 10.1016/j.ijhcs.2006.06.009
- Tuch, A. N., Bargas-Avila, J. a. & Opwis, K. (2010). Symmetry and aesthetics in website design: It's a man's business. *Computers in Human Behavior*, 26(6), 1831–1837. doi:10.1016/j.chb.2010.07.016
- Tuch, A. N., Presslauer, E. E., Stöcklin, M., Opwis, K. & Bargas-Avila, J. A. (2012a). The role of visual complexity and prototypicality regarding first impression of websites: Working towards understanding aesthetic judgments. *International Journal of Human-Computer Studies*, 70, 794–811. doi: 10.1016/j.ijhcs.2012.06.003



- Tuch, A. N., Roth, S. P., Hornbæk, K., Opwis, K. & Bargas-Avila, J. A. (2012b). Is beautiful really usable? Toward understanding the relation between usability, aesthetics, and affect in HCI. *Computers in Human Behavior*, 28(5), 1596–1607. doi: 10.1016/j.chb.2012.03.024
- von Ehrenfels, C. (1890). Über Gestaltqualitäten. *Vierteljahresschrift für wissenschaftliche Philosophie*, 14, 249-292.
- Wang, X., Huang, Y., Ma, Q. & Li, N. (2012). Event-related potential P2 correlates of implicit aesthetic experience. *NeuroReport*, 23, 862–866. doi:10.1097/WNR.0b013e3283587161
- Winston, A. S., & Cupchik, G. C. (1992). The evaluation of high art and popular art by naive and experienced viewers. *Visual Arts Research*, 18, 1-14. <http://www.jstor.org/stable/20715763>
- Xing, J. & Manning, C. A. (2005). Complexity and automation displays of air traffic control: Literature review and analysis. DTIC Document. Retrieved from <http://oai.dtic.mil/oai/oai?verb=getRecord&metadataPrefix=html&identifier=ADA460107>
- Zeki, S. (2001). Artistic creativity and the brain. *Science*, 293(5527), 51-52. doi: 10.1126/science.1062331

**Figure Captions**

Figure 1: EEG-signal as a function of expertise, aesthetics and anterior – posterior regions of interests.

Figure 2: EEG-signal as a function of expertise, aesthetics and left-right regions of interests.

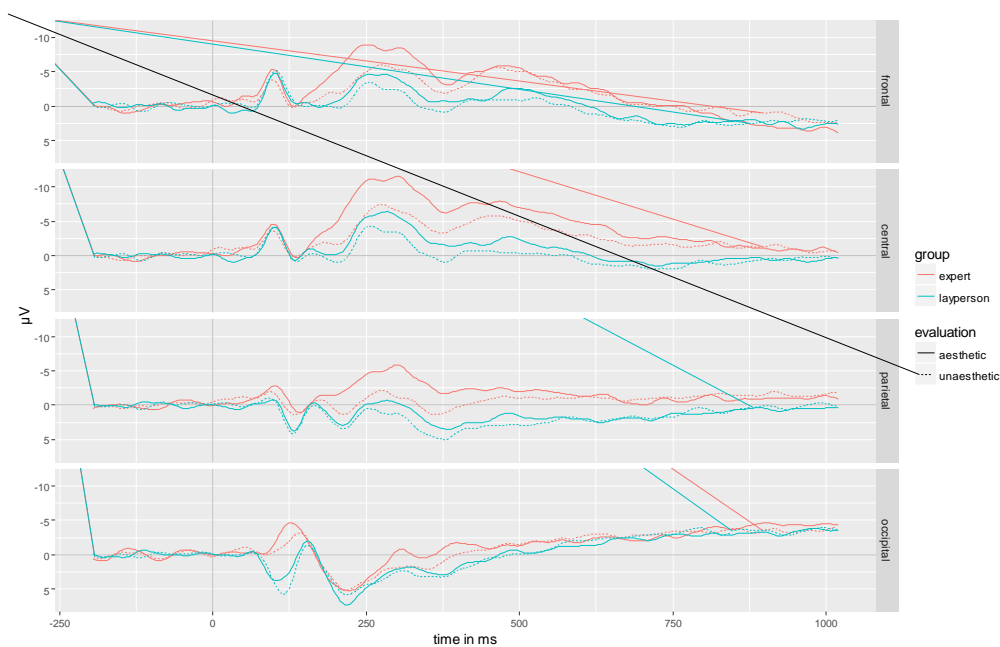


Figure 1

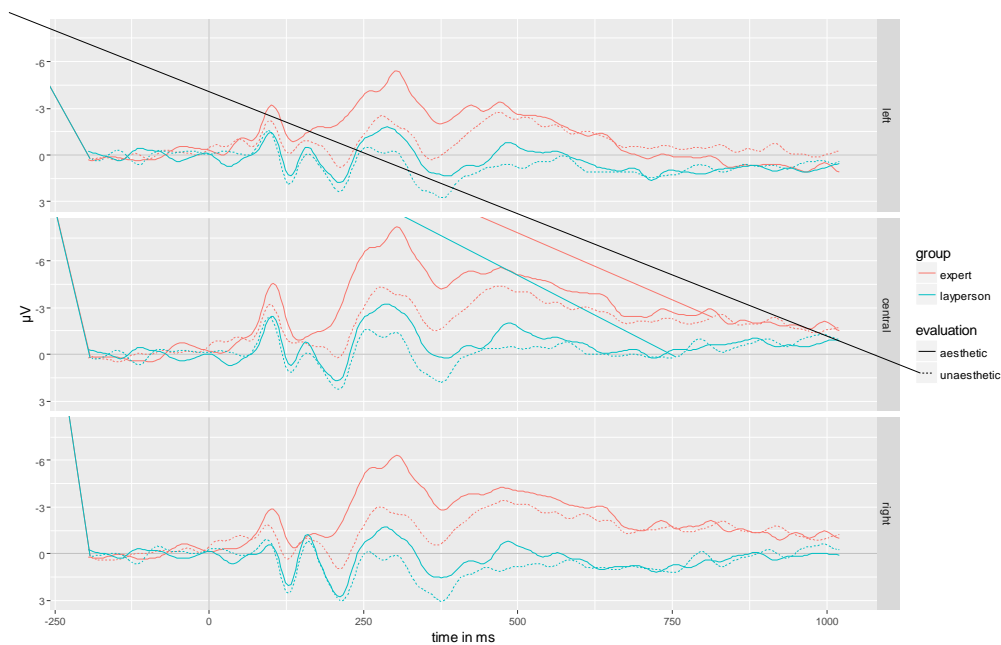


Figure 2

### Footnotes

<sup>1</sup> We are aware that visual complexity measured in bytes is a crude measure of complexity. It is a numerical measure ignoring functional and practical aspects as well as the integration by the observer (see Xing and Manning, 2005, for a review of definitions of complexity).