

EFFECTS OF INTERFACE AESTHETICS ON TEAM PERFORMANCE IN A VIRTUAL TASK

Research in Progress

Maria Douneva; University of Muenster, Germany, maria.douneva@uni-muenster.de

Russell Haines; Old Dominion University, USA, rhaines@odu.edu

Meinald T. Thielsch; University of Muenster, Germany, thielsch@uni-muenster.de

Abstract

HCI research has acknowledged the importance of aesthetics for user interfaces by examining its effects on users' attitudes and reactions. However, evidence for effects on task performance is mixed. By manipulating chat background colour in a within-subjects design, this study investigates the effects of a less attractive vs. attractive chat interface in a collaborative setting. Participants performed the simulated role of a nurse, a doctor, laboratory technician, or specialist as a member of an emergency response team that had to diagnose patients within a given time. The data of 184 participants during three rounds of the same task was analysed for effects of the colour manipulation on emotion, affect, and team performance. Chat background colour was randomly varied in round 2 and 3. Although participants clearly preferred the attractive to the less attractive version, analyses of variance revealed that neither their attitudes nor their performance as a group was significantly influenced by the colour manipulation. We discuss (1) further analyses on the individual level and qualitative analyses of the chat protocols, (2) possible explanations such as the performance criterion, nature of the task, and aesthetics manipulation, and (3) future directions for research on the link between aesthetics and performance.

Keywords: team performance, collaboration, aesthetics, interface design

1 Introduction

The notion that the appearance of a technological artefact influences the decision performance of an information system user dates back to the foundational Minnesota experiments (Dickson, Senn, and Chervany, 1977). In the ensuing years, the notions of mode of presentation were largely discarded from information systems research in favour of an emphasis on increasing the ease of use, usefulness and/or “playfulness” (i.e., experiential enjoyment) of the technology (Moon and Kim, 2001). Over the past 15 years, however, research showed repeatedly that users' needs go beyond pure usability in the design of graphical user interfaces, leading to a shift in focus to a more experiential perspective in HCI. This present perspective includes user perceptions of facets such as visual aesthetics (e.g., Bargas-Avila and Hornbæk, 2011; Hassenzahl and Tractinsky, 2006).

Since the beginning of aesthetics research in HCI in the mid-nineties, various studies have demonstrated the importance and impact of aesthetics on perceived usability, user satisfaction, or users' reactions such as preference, urge to buy impulsively, or intention to revisit (for an overview, see Moshagen and Thielsch, 2010, p. 691). Additionally, it is well known that aesthetic responses occur immediately at first sight of a technological artefact (cf. Thielsch and Hirschfeld, 2012; Tuch et al., 2012). Unfortunately, while one can find several effects of aesthetics on attitudes and user reaction measurements, only a few studies have directly examined the effects of aesthetics on performance. Some stud-

ies indicate that visual aesthetics may be capable of improving individual performance (e.g., Miller, 2011; Moshagen, Musch, and Göritz, 2009; Sonderegger and Sauer, 2010; Reppa and McDougall, 2015), while others did not find such effects (e.g., Ilmberger, Schrepp, and Held, 2009; Schmidt, Liu, and Sridharan, 2009; Thielsch, Nikolaeva, and Kebeck, 2009; van Schaik and Ling, 2009).

What might be the reason behind these mixed results? First, aesthetics is a multidimensional construct influenced by a number of factors, including colour, visual complexity, symmetry and balance, familiarity, and many more (for an overview, see Reppa and McDougall, 2015). If these factors are not separated from each other, they might act as confounding variables. For example, manipulations of appeal can also affect visual complexity (e.g., Sonderegger and Sauer, 2010). Second, performance is task-dependent, so the influence of aesthetics on performance may similarly depend on the task at hand. Prior research suggests that aesthetics facilitates performance mainly in problem-solving situations, and more when the task is difficult than when the task is easy (Moshagen, Musch, and Göritz, 2009; Reppa and McDougall, 2015).

Since the process of getting a team organised for work is a relatively difficult task, aesthetics should affect performance in a collaborative environment. The purpose of this research is to investigate whether aesthetics, manipulated through colour, impacts performance in a collaborative setting by comparing individuals that use an aesthetically attractive communication tool with users of a less attractive version of the same tool. To our knowledge, this is the first study that investigates the link between aesthetics and performance in a non-creative collaboration task.

2 Theoretical Background

The term “aesthetics” is not consistently used in the literature –some designations for the same construct use “beauty” or “pleasure”. Moshagen and Thielsch (2010, p. 690) provided a definition which describes website aesthetics as “an immediate pleasurable subjective experience that is directed toward an object and not mediated by intervening reasoning”. As a component of user experience, a user’s judgment of the aesthetics of a technology tool can be made in less than five seconds (van Schaik and Ling, 2009).

Aesthetics has been linked with both perceived and actual usability of a technology tool, leading to the contention that “What is beautiful is usable.” (Tractinsky, Katz, and Ikar, 2000). However, subsequent research suggests that this only applies to perceived usability and that usability and aesthetics are separately judged (van Schaik and Ling, 2009). These equivocal results have led researchers examining the link between aesthetics and usability to propose that additional variables mediate the relationship between technology tool aesthetics and usability (cf. Lee and Koubek, 2012).

Norman (2002) argues that aesthetics influences the cognitive system and thus problem solving by changing the emotional state. The results of Moshagen and colleagues (2009) were in line with this *positive affect mediation hypothesis*, but they did not directly test affect mediation or control for possibly confounding variables. Another possible account is the *prolongation of joyful experience/increased motivation hypothesis* (Sonderegger and Sauer, 2010). While the first part of the hypothesis explains decrements in performance by users seeking to prolong enjoyment with aesthetically appealing products instead of solving the task, the second part states that an enhancement in performance might result from being in a flow situation when working in aesthetically pleasing environments. Van Schaik and Ling (2009) argue that context and experience impact the relationship between aesthetics and usability. In our study, we chose a collaboration context where we can test the mediating effect of emotional state by including both emotional and performance variables.

Teams that need to accomplish a collaborative task actually have several subtasks to accomplish before the actual task completion: inception, problem solving, and conflict resolution (McGrath, 1991). In our experimental context, the first three stages were separated from the actual task accomplishment period by giving teams a distinct time in which to get organised. The aesthetic experimental manipula-

tion was done during this organisation period, meaning individuals used one of two communication tools – one aesthetically less attractive than the other.

3 Hypotheses

From a systems perspective, we wanted to learn the extent to which a designer's aesthetic missteps might have implications for the performance of teams in a collaborative setting. In an aesthetically pleasing environment, we expect users to have a more positive affect in general and specifically toward their team, therefore they would be more likely to engage with their team, eventually resulting in higher team performance:

H1: Users working in an aesthetically pleasing environment have more positive emotions toward their team and their affect is more positive.

H2: Team performance is enhanced directly through team members having been exposed to an aesthetically pleasing environment, or indirectly through team members having a more positive affect after being exposed to an aesthetically pleasing environment.

In our experiment, we wanted to test whether these effects would hold in a more “natural” collaborative environment. Thus, we wanted to examine the effect of aesthetics separately from the effect of the novelty of a web page or tool. For this reason, we allowed our participants to get used to their own role in a team and to familiarise themselves with the communication tool in what amounted to a trial or training team. After they were trained, they used a new version of the communication tool, one of which was aesthetically less pleasing than the other.

4 Experimental method

4.1 Participants

Participants in the experiment were students enrolled in a general education information literacy course at a public U.S. university. The data of 184 students (53.3% male, mean age = 21 years, $SD = 4.8$) were included¹ in this study. The experimental task was part of a class exercise on team collaboration. All interaction in the setting was via computer-mediated communication: team members communicated via chat and email using a browser-based application that was custom made for this study. Experimental sessions were conducted in a computer classroom in which each participant was seated in front of a computer.

4.2 Experimental setup and treatments

The experimental manipulation used for this paper was a chat screen colour treatment: attractive vs. less attractive (within-subjects design). Every participant performed the simulated role of a nurse, a doctor, laboratory technician, or specialist as a member of an emergency response team. The assignment to one of these four possible roles² took place at the beginning of the experiment. Participants completed three rounds of examining and diagnosing patients as a team. Their role remained the same,

¹ The participants were asked, “Are you color blind or do you have any vision disorders that affect your ability to use a computer?” as part of the demographic questionnaire. Eliminating those that indicated “Yes” ($n = 4$) had no effect on the results.

² When the class size was a multiple of three, the team had only three members where the doctor made the final diagnosis. All analyses were conducted with all teams ($N = 184$) and with only three-member teams ($N = 144$). Results are displayed for all teams because there were no changes in the results due to team size.

but they were randomly assigned to a new team after the first and second diagnosing periods, meaning that participants diagnosed patients with three different teams.

The chat room colour design randomly changed between the rounds (see the next section for a detailed description). The colour manipulation consisted of either an attractive, dark blue background (html colour #59A2CF) or a less attractive, greenish blue background (html colour #00EEBB, see Figure 1) during the chat period. This colour manipulation was shown in prior studies on aesthetics to lead users to differ significantly in their perceptions of attractiveness (e.g. study 6 in Moshagen and Thielsch, 2010).

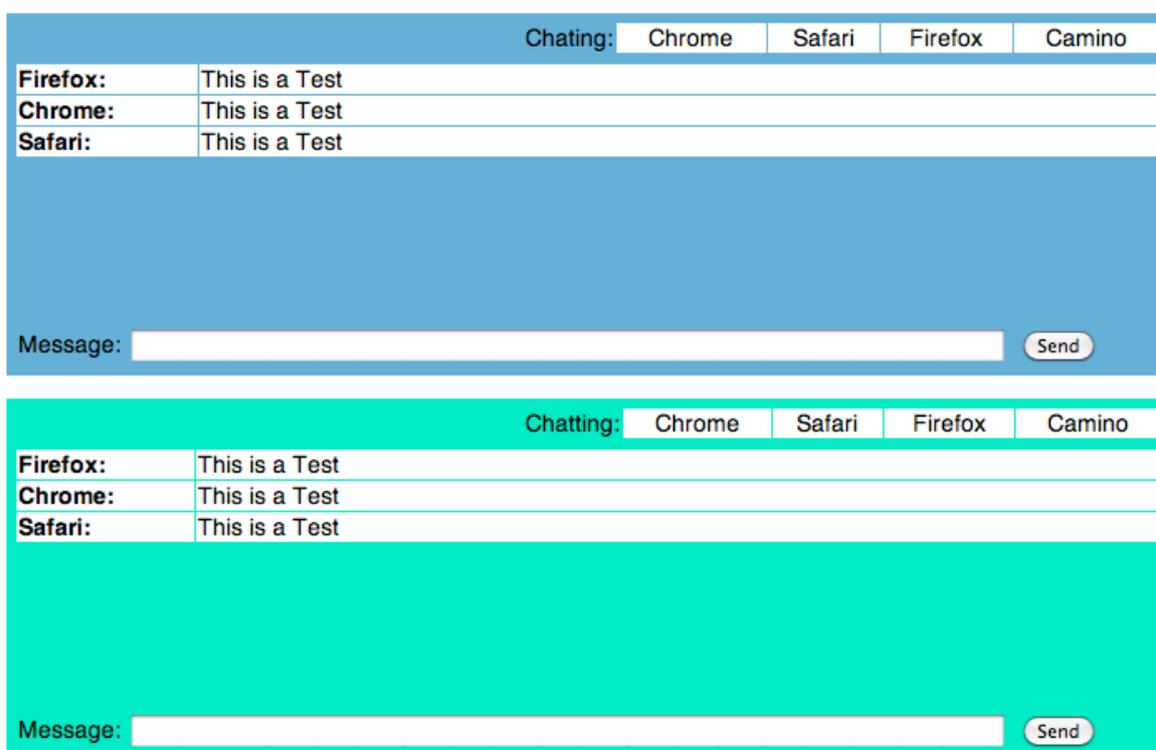


Figure 1. Screenshots of both the attractive (top) and less attractive (bottom) chat background

4.3 Procedure

In all sessions, the same administrator read the same, scripted instructions, which were also displayed on the participants’ screens. Participants first received informed consent information and completed a demographic questionnaire. Next, the participants were assigned to a job and read job-specific instructions. Finally, the participants were divided into teams and met in a chat room. The first chat session was in a gold-coloured (#d3d88f) chat room² for all participants and lasted five minutes to give the participants time to get to know each other and to discuss how to complete a task with which they were unfamiliar. Later chat sessions were three minutes long and varied in their colour scheme by participant. After chatting, the participants completed a post-chat questionnaire (see Measures section).

During each diagnosing period, each team member was shown a list of the ten patients that needed to be diagnosed. The jobs were organised such that the final diagnosis could not be correctly made without a laboratory test (performed by the laboratory technician), patient examination (performed by the

² The basic colour design was inspired by a real software application and should serve as “baseline condition”.

doctor), and patient interview (performed by the nurse). Furthermore, the patient examination requested information from the patient interview, and a specific laboratory test would be called for after the patient examination was finished. Thus, a four-step process was encouraged (Figure 2):

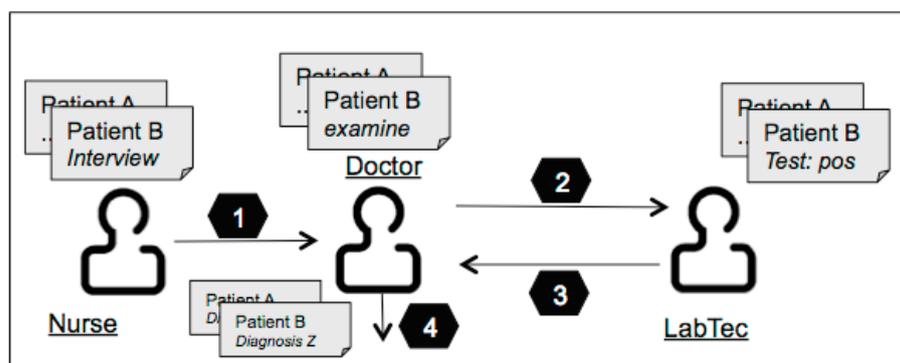


Figure 2. Experimental setup depicting a three person team and the information flow necessary for diagnosing patients. In a four person team, a specialist made the diagnosis.

First, the nurse would conduct a patient interview and pass that information to the doctor. Next, the doctor would enter the results of the patient interview, conduct the patient examination, and request the required test from the laboratory technician. Third, the laboratory technician would perform the laboratory test requested by the doctor and send the results back to the doctor. Finally, the doctor or specialist would determine the final diagnosis for the patient by entering the results of the laboratory test. When the team had four members, the doctor and laboratory technician sent their information to the specialist, who then performed the final diagnosis.

After the first and second diagnosing periods, participants were divided into new teams and chatted with their new team about how they would diagnose patients. This enables the teams to experience “swift trust” (Meyerso, Weick, and Kramer, 1996) and is more environmentally valid than the typical single shot teams used in most collaboration research (Pinsonneault and Heppel, 1997).

The design of the task allowed patients to be interviewed, examined, or tested in any order. Thus, it was possible for the team members to behave rather chaotically, picking patients to process at random. However, there was a time penalty for performing an examination without knowing the results of the patient interview and for choosing the wrong test to perform.

4.4 Measures

The questionnaire items used for this study were gathered after the chat period and at the conclusion of the experiment. The questions for the scales were intermingled and displayed in a randomised order for each participant. All questions were 7-point Likert-type and anchored *Strongly Disagree – Strongly Agree*. After each chat period and before the diagnosing period, the following measures were administered:

Emotion toward team is defined as the extent to which team members experience a state of high energy and pleasurable engagement toward working with their team, and was assessed with four questions (Cronbach’s $\alpha = .95$) specifically related to emotions regarding teamwork: (1) I have a positive attitude toward working with this team, (2) I am in a good mood about working with the other members of this team, (3) I think being a member of this team will be enjoyable for me, (4) I expect to have a nice time working with this team. Participants answered these questions after chatting with their team and before they start diagnosing patients with this same team.

Affect is defined as the extent to which one feels calm, peaceful, and untroubled, and was measured with the “serenity” subscale of the PANAS (Positive Affect Negative Affect Scale; Watson, Clark, and

Tellegan, 1988): (1) During the chat discussion, I felt relaxed, (2) During the chat discussion, I felt at ease, (3) During the chat discussion, I felt calm.

Colourfulness is the extent to which the colours of a web site are perceived to be aesthetically pleasing, and served as manipulation check for the chat screen colour treatment. The questions were derived from the corresponding subscale of the VisAWI (Visual Aesthetics of Websites Inventory; Moshagen and Thielsch, 2010; $\alpha = .89$ and satisfactory convergent, divergent, discriminative, and concurrent validity): (1) The colour composition of the chat screen was attractive, (2) The colours in the chat room did not match, (3) The choice of colours for the chat room was botched, (4) The colours in the chat room were appealing.

At the very end of the experiment, participants were asked to rate the colour scheme of both the attractive and the less attractive chat background. Responses to these items were on a 7-point scale, anchored *Strongly Dislike – Strongly Like*.

Team performance as the main dependent variable was measured as number of diagnosed patients after each round, meaning that it could possibly range from 0 to 10.

5 Results

In round 1, there was no colour manipulation in order to establish a baseline condition and to avoid effects of the tools novelty (all participants saw a gold background). Accordingly, only data from rounds 2 and 3 were analysed. Consistent with our past experience, the first round was rather confusing and stressful for the participants. After completing the first round, the task and the participants' role were quite clear and collaborative practices had evolved. As such, rounds two and three more closely approximated a real-life work arrangement: members of the teams had a good idea about how to perform their roles and their tasks, and they had worked in teams before and knew how collaboration could be organised.

5.1 Colour manipulation check

Direct ratings of colour schemes at the end of the experiment were significantly different from each other: Participants indicated that they liked the attractive scheme ($M = 5.36$, $SD = 1.66$) more than the less attractive scheme ($M = 4.45$, $SD = 1.98$; $t(165) = -4.36$, $p < .001$), which represents a medium-sized effect of $d = 0.5$. However, when comparing participants who saw the attractive vs. the less attractive colour scheme during the chat period, their evaluations of *colourfulness* immediately after the chat discussions did not significantly differ from each other (attractive: $M = 4.72$, $SD = 1.32$ vs. less attractive: $M = 4.51$, $SD = 1.28$; $t(182) = -1.11$, $p = .27$, $d = 0.16$).

5.2 Effects of the colour manipulation on emotion toward team and affect

In round 2, the colour manipulation was implemented for the first time. Using Pillai's trace in a multivariate analysis of variance (MANOVA), the results indicate that this manipulation did not have a significant effect on either emotion toward the team or affect, $V = 0.27$, $F(2, 181) = 2.495$, $p = .09$.

Participants who saw the attractive chat background ($N = 92$) indicated a better emotion toward the team ($M = 5.47$, $SD = 1.46$) than participants who saw the less attractive chat background ($N = 92$, $M = 5.19$, $SD = 1.62$), but this difference is not significant ($p = .22$, $d = 0.18$). The ratings for affect did not differ between the two conditions (attractive: $M = 5.21$, $SD = 1.52$; less attractive: $M = 5.20$, $SD = 1.64$, $p = .99$, $d = 0.01$).

In round 3, participants were again randomly assigned to one of the two colour conditions, resulting in four groups ($N = 46$ each): two that saw the same colour scheme again and two that saw the other colour scheme. Using Pillai's trace in a MANOVA, the results indicate that there were no significant differences between these four groups regarding emotion toward the team ($p = .26$) or affect ($p = .14$), $V = 0.34$, $F(6,360) = 1.04$, $p = .39$.

Descriptive statistics are provided in table 1, indicating that participants who remained in the less attractive condition had the lowest emotion and affect scores. Still, none of the Bonferroni post-hoc tests showed significant differences between the four conditions. Overall, our H1 cannot be confirmed.

Colour scheme in round 2/3	Emotion toward team (M, SD)	Affect (M, SD)
Attractive/attractive	5.51 (1.59)	5.40 (1.59)
Attractive/less attractive	5.61 (1.46)	5.70 (1.42)
Less attractive/attractive	5.64 (1.50)	5.60 (1.64)
Less attractive/less attractive	5.09 (1.52)	5.00 (1.44)

Table 1. Means and standard deviations of emotion toward team and affect scales after round 3 for all four colour treatment groups.

5.3 Effects of the colour manipulation on team performance

Across all three rounds, team performance ranged from 0 to 6 with a mean of 1.87 ($SD = 1.57$). Analysis of variance (ANOVA) revealed that performance significantly differed by round ($F(2, 549) = 46.79, p < .001$) with being lowest in round 1 ($M = 1.05, SD = 1.0$), higher in round 2 ($M = 2.07, SD = 1.62$) and highest in round 3 ($M = 2.48, SD = 1.65$), which represents a large effect of $\omega^2 = .14$ (between-groups effect divided through total amount of variance and adjusted for estimation of the effect size in the population).

ANOVA was also used to investigate the effects of the colour manipulation in rounds 2 and 3. In both rounds, being in the less attractive vs. attractive colour condition did not alter team performance significantly (round 2: $F(1, 182) = 0.30, p = .59$, round 3: $F(1, 182) = 0.65, p = .42$). Our H2 is thus not confirmed. Still, there was a tendency for better performance after seeing the attractive colour in round 2 (2.13 vs. 2.00, $d = 0.08$) and in round 3 (2.58 vs. 2.38, $d = 0.12$).

6 Discussion

6.1 Main findings and outlook

We found that the use of a less aesthetically pleasing communication tool had, in general, no significant effect on the performance of collaborative team members who are familiar with the system – neither directly nor indirectly through a change in attitudes or perceptions. Thus, we suggest that designers of communication tools for a collaborative working context should focus more on usability in general than on aesthetics in particular. Our participants clearly preferred one of the communication tools over the other when they had to directly rate the two versions. However, when one or the other of the two was forced on them, it made little difference in their rating of the colour scheme, their emotion toward the team, their affect, or the performance of their team.

We note that our participants were relatively practiced team members at the point when they saw the aesthetically varied communication tools, and their only task during the chat discussion was to ensure that their new team was organised and ready to work. We caution that in a context where team members are just becoming familiar with a team and task, a new, less-aesthetically pleasing communication tool might be the so-called “straw that breaks the camel’s back,” and lead to more serious team or individual problems (cf. Thielsch, Blotenberg, and Jaron, 2014). More research could examine the extent to which aesthetics affects its initial use and therefore the initial adoption decision of a technology in a team context.

Our users may have also been more in a “flow” situation and not paying particular attention to differences in the aesthetics of the technology tool they were using. However, the post-experiment manipu-

lation check was positive and we did observe a few mentions of the different appearances in the chat room discussions, which suggests that the participants at least noticed the difference, even though it did not seem to affect their emotion toward their team, affect, or performance. Additionally, aesthetics might have influenced individual performance in a way that would not show up in the group analyses. Thus, in a second step we are going to take a look at the individual data to examine whether the tools led to differences in individual performance (e.g., number of interviewed patients for the nurse) or the number and quality of comments that they made. As there were clear gains in performance from round to round, these individual analyses might offer more insight into factors that influenced perceptions and performance.

6.2 Limitations and future research

There are several possible explanations why we, in contrast to some other studies, did not find significant effects of interface aesthetics on performance.

First, our performance measure was the number of solved cases within a restricted amount of time. Effects may have occurred with a different operationalisation, for example completion time (Moshagen et al., 2009), or with another type of task. Following the line of Norman's (2002) argumentation, attractiveness leads to a better mood which in turn encourages creative thinking. Participants in our study could not be creative while solving the task; hence, presenting participants with a problem that requires a more creative approach may lead to different results, as was the case in Bhagwatwar and colleagues' study (2013) where visual priming led to enhanced team performance in a brainstorming task.

Second, manipulating the colourfulness alone may not be strong enough to elicit substantial effects. It is also only one way of changing perceived aesthetics. Although colourfulness was repeatedly found to be a crucial characteristic of web interfaces (e.g., Lindgaard, 2007; Cyr, Head, and Larios, 2010; Peak et al., 2014), one could instead or additionally manipulate aspects such as diversity (see Moshagen and Thielsch, 2010) or visual complexity (Tuch et al., 2012). Still, one has to be cautious that a strong aesthetics manipulation is not perceived as unrealistic. With our colour manipulation, we aimed for a realistic design with the use of colours a designer might build into a website.

Third, although preferences for web interfaces can form after milliseconds (e.g., Lindgaard, 2006; Thielsch and Hirschfeld, 2012), this primarily happens when no context is provided, that means when a less attractive and attractive version are just presented and then rated compared to giving participants tasks or instructions and thus providing a context (van Schaik and Ling, 2009). This corresponds to our findings: Participants had a clear preference when presented with the two chat interface versions without any tasks besides rating the attractiveness, whereas this preference did not manifest itself when participants were working on their task while being exposed to the interfaces. In general, perceptions of aesthetics seem to be context-dependent: In a less serious context, a halo effect of aesthetics is more probable and can even compensate for low usability (de Angeli, Sutcliffe, and Hartmann, 2006).

Lastly, the importance of context also applies to the user side: Perceptions of aesthetics and usability depend on users' cultural background (e.g., Simon, 2001; Gevorgyan and Manucharova, 2009) and for example lead to better performance on culturally adapted interfaces (Reinecke and Bernstein, 2011). Accordingly, future studies should take contextual factors on the user side into account when deriving conclusions about the relationship between aesthetics, attitudes, perceptions, and performance.

In conclusion, from our point of view, research on performance effects of aesthetics should be embedded in a framework, ordering task and context characteristics (e.g., simple vs. complex, modes of use, cultural context), time and focus of evaluation (e.g., first vs. overall impression, individual vs. team focus) as well as design variables (e.g., facets of aesthetics, low vs. high usability settings). We hope that our results (and further analyses on the individual level) will contribute to such a framework and better understanding of the processes behind the currently mixed results of aesthetics on performance.

References

- Ayyagari, A., Grover, V., and R. Purvis (2011). "Technostress: Technological Antecedents and Implications." *MIS Quarterly* 35 (4), 831-858.
- Bargas-Avila, J. A. and K. Hornbæk (2011). "Old wine in new bottles or novel challenges: a critical analysis of empirical studies of user experience." *CHI'11 Proceedings of the 2011 annual conference on Human factors in computing systems*, 2689-2698.
- Bhagwatwar, A., Massey, A. and A. R. Dennis (2013). "Creative Virtual Environments: Effects of Supraliminal Priming on Team Brainstorming". *46th Hawaii International Conference on System Sciences (HICSS)*, Maui, Hawaii, January 7-10.
- Cyr, D., Head, M., and H. Larios (2010). "Colour appeal in website design within and across cultures: A multi-method evaluation." *International Journal of Human-Computer Studies* 68, 1-21.
- De Angeli, A., Sutcliffe, A., and J. Hartmann (2006). "Interaction, usability and aesthetics: What influences users' preferences?" *DIS 2006 Conference Proceedings*, 271-280.
- Dickson, G. W., Senn, J. A., and N. L. Chervany (1977). "Research in management information systems: The Minnesota experiments." *Management Science* 23 (9), 913-923.
- Fritz, C. and S. Sonnentag (2009). "Antecedents of Day-Level Proactive Behavior: A Look at Job Stressors and Positive Affect During the Workday." *Journal of Management* 35 (1), 94-111.
- Gevorgyan, G. and N. Manucharova (2009). "Does culturally adapted online communication work? A study of american and chinese internet users' attitudes and preferences toward culturally customized web design elements." *Journal of Computer-Mediated Communication* 14 (2), 393-413.
- Hassenzahl, M. and N. Tractinsky (2006). "User experience - a research agenda." *Behaviour & Information Technology* 25, 91-97.
- Ilmberger, W., Schrepp, M., and T. Held (2009). "Was verursacht den Zusammenhang zwischen Ästhetik und Usability?" In: *Mensch & Computer 2009*. Ed. by H. Wandke, S. Kain, and D. Struve. München: Oldenbourg, pp. 383-392.
- Kerr, N. L. and G. Hertel (2011). "The Köhler Group Motivation Gain: How to Motivate the 'Weak Links' in a Group." *Social and Personality Psychology Compass* 5 (1), 43-55.
- Lee, S. and R. J. Koubek (2012). "Users' perceptions of usability and aesthetics as criteria of pre- and post-use preferences." *European Journal of Industrial Engineering* 6 (1), 87-117.
- Lindgaard, G., Fernandes, G., Dudek, C., and J. Brown (2006). "Attention web designers: You have 50 milliseconds to make a good first impression!" *Behavior & Information Technology* 25 (2), 115-126.
- Lindgaard, G. (2007). "Aesthetics, visual appeal, usability and user satisfaction: What do the user's eyes tell the user's brain?" *Australian Journal of Emerging Technologies and Society* 5(1), 1-14.
- McGrath, J. E. (1991). "Time, interaction, and performance (TIP) A Theory of Groups." *Small Group Research* 22(2), 147-174.
- Meyerson, D., Weick, K. E., and R. M. Kramer (1996). "Swift Trust and Temporary Groups." In: *Trust in Organizations: Frontiers of Theory and Research*. Ed. by R. M. Kramer and T. R. Tyler. Thousand Oaks, CA: SAGE Publications, pp. 166-195.
- Miller, C. (2011). "Aesthetics and e-assessment: the interplay of emotional design and learner performance." *Distance Education* 32 (3), 307-337.
- Moon, J. W. and Kim, Y. G. (2001). "Extending the TAM for a World-Wide-Web context." *Information & Management* 38 (4), 217-230.
- Moore, J. E. (2000). "One road to turnover: an examination of work exhaustion in technology professionals." *MIS Quarterly* 24 (1), 141-168.
- Moshagen, M., Musch, J., and A. S. Göritz (2009). "A blessing, not a curse: Experimental evidence for beneficial effects of visual aesthetics on performance." *Ergonomics* 52, 1311-1320.
- Moshagen, M. and M. T. Thielsch (2010). "Facets of visual aesthetics." *International Journal of Human-Computer Studies* 68, 689-709.

- Norman, D. A. (2002). "Emotion and design: Attractive things work better." *Interactions Magazine* ix (4), 36-42.
- Peak, D., Prybutok, V., Gibson, M., Wu, Y., and C. Xu (2014). "User perceptions of aesthetic visual design variables within the informing environment: A web-based experiment." *Informing Science: the International Journal of an Emerging Transdiscipline* 17, 25-57.
- Pinsonneault, A. and N. Heppel (1997). "Anonymity in group support systems research: A new conceptualization, measure, and contingency framework." *Journal of Management Information Systems* 14 (3), 89-108.
- Reinecke, K. and A. Bernstein (2011). "Improving performance, perceived usability, and aesthetics with culturally adaptive user interfaces." *ACM Transactions on Computer-Human Interaction (TOCHI)* 18 (2), 8.
- Reppa, I. and S. McDougall (2015). "When the going gets tough the beautiful get going: aesthetic appeal facilitates task performance." *Psychonomic Bulletin & Review* doi:10.3758/s13423-014-0794-z
- Schmidt, K. E., Liu, Y., and S. Sridharan (2009). "Webpage aesthetics, performance, and usability: Design variables and their effects." *Ergonomics* 52 (6), 631-643.
- Simon, S. J. (2001.) "The impact of culture and gender on web sites: An empirical study." *The DATA BASE for Advances in Information Systems* 32 (1), 18-37.
- Sonderegger, A. and J. Sauer (2010). "The influence of design aesthetics in usability testing: Effects on user performance and perceived usability." *Applied Ergonomics* 41 (3), 403-410.
- Thielsch, M. T., Blotenberg, I. and R. Jaron (2014). "User evaluation of websites: From first impression to recommendation." *Interacting with Computers* 26 (1), 89-102.
- Thielsch, M.T. and G. Hirschfeld (2012). "Spatial frequencies in aesthetic website evaluations – explaining how ultra-rapid evaluations are formed." *Ergonomics* 55 (7), 731-742.
- Thielsch, M. T., Nikolaeva, D., and G. Kebeck (2009). "Beeinflusst Ästhetik die Performanz bei Suchaufgaben im Web?" In: *Mensch & Computer 2009*. Ed. by H. Wandke, S. Kain, and D. Stroove. München: Oldenbourg, pp. 487-490.
- Tractinsky, N., Katz, A. S., and D. Ikar (2000). "What is beautiful is usable." *Interacting with Computers* 13 (2), 127 - 145.
- Tuch, A. N., Presslauer, E. E., Stöcklin, M., Opwis, K., and J. A. Bargas-Avila (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 Schaik, P. and J. Ling (2009). "The role of context in perceptions of the aesthetics of web pages over time." *International Journal of Human-Computer Studies* 67, 79-89.
- Watson, D., Clark, L. A., and A. Tellegan (1988). "Development and validation of brief measures of positive and negative affect: The PANAS scales." *Journal of Personality and Social Psychology* 54 (6), 1063-1070.