Computers in Human Behavior 44 (2015) 81–95

Contents lists available at ScienceDirect

Computers in Human Behavior

journal homepage: www.elsevier.com/locate/comphumbeh

Emotional design in multimedia learning: Differentiation on relevant design features and their effects on emotions and learning

Steffi Heidig [a,*], Julia Müller [a], Maria Reichelt [b]

[a] University of Erfurt, Psychology Department, P.O. Box 900 221, 99105 Erfurt, Germany
[b] Bauhaus-University Weimar, Instructional Design, Geschwister-Scholl-Straße 7, 799423 Weimar, Germany

ARTICLE INFO

Article history:
Available online 5 December 2014

Keywords:
Multimedia learning
Emotional design
Emotion
Aesthetics
Usability

ABSTRACT

Previous research into multimedia learning has mainly focused on cognitive factors to investigate different instructional conditions and design principles. Emotional factors have so far been widely neglected. However, recent studies showed that the emotional design of multimedia learning material can evoke positive emotions in learners that in turn facilitate the learning process. Following this lead, our study aims to further explore the potential of an emotional design. We seek to differentiate the current findings by systematically deducing emotionally relevant design features and also taking into account negative emotional states. In order to deduce relevant design features, we adopt concepts from web design. German college students ($N = 334$) were assigned to one of nine conditions, created by two design factors (classical vs. expressive aesthetics), each with two levels (high vs. low) and a usability factor (high vs. low usability) as well as a control group (no color/gray scale). Unexpectedly, objective differences in aesthetics or usability did not affect learners' emotional states. However, the perceived aesthetics and usability positively affected the emotional states of the learners. Learners' emotional states had a minor impact on learning outcomes but a larger impact on learners' intrinsic motivation, including the motivation to continue working with the material.

Instructional conditions to enhance computer-based multimedia learning have been tested primarily with respect to cognitive factors (e.g., Mayer, 2009; Mayer & Moreno, 2003; Plass, Moreno, & Brünken, 2010). Emotional factors have widely been neglected in multimedia learning research so far, despite the existence of several unanswered questions (e.g., Leutner, 2014; Park, Plass, & Brünken, 2014). How can we design multimedia learning materials that are appealing while still being effective for learning? Are design appeal and learning effectiveness in conflict with each other? Or is the opposite true, and a felicitous design can be advantageous for learning? The integration of appealing but interesting design elements that are superfluous to learning has been rather critically discussed in multimedia learning research. Although these additional elements are assumed to positively affect learners' emotions and motivation, they are believed to induce an extraneous cognitive load and therefore harm learning (seductive details: e.g., Garner, Gillingham, & White, 1989; Harp & Mayer, 1997; Lehman, Shraw, McCrudden, & Hartley, 2007; Lenzner, 2009; Mayer, Heiser, & Lonn, 2001; Rey, 2012; coherence principle: e.g., Mayer, 2009). Recently, an *emotional design* for multimedia learning materials has been proposed (Plass, Heidig, Hayward, Homer, & Um, 2014; Um, Plass, Hayward, & Homer, 2012). Emotional design is the use of visual design elements in multimedia learning that can evoke positive emotions and therefore facilitate learning. An emotional design for multimedia learning materials does not necessarily require additional elements, but may instead change intrinsic design elements such as color, layout or round vs. square shapes. In line with this assumption, Um et al. (2012) and Plass et al. (2014) showed that warm colors and round shapes may evoke positive emotions in learners that in turn facilitate motivation and learning outcomes. However, we are only beginning to understand the role emotions may play in multimedia learning. The current study ties in with this new line of research and aims to further explore the potentials of an emotional design in multimedia learning. To this end, this study examines (1) whether intrinsic design features of multimedia learning material affect learners' emotional states and (2) how emotional states that are experienced during learning may affect the learning process. Whereas previous studies (Plass et al., 2014; Um et al., 2012) provided initial insights into the role of selected design features (warm colors, round shapes) and positive emotions, we seek to broaden the picture by systematically deducing which design features to manipulate, by investigating the effect of negative

* Corresponding author. Tel.: +49 361 7372750.
E-mail address: steffi.heidig@uni-erfurt.de (S. Heidig).

http://dx.doi.org/10.1016/j.chb.2014.11.009
0747-5632/

emotional states on learning rather than looking at positive emotional states alone. In order to deduce emotionally relevant design features, we take a multidisciplinary view by adopting concepts from web design.

1. Deducing emotional design features

As with developments in the field of multimedia learning, the main focus in the field of human–computer interaction in general and web design in particular has primarily been on cognitive and pragmatic issues, namely system usability (e.g., Hassenzahl & Tractinsky, 2006; Moshagen, Musch, & Göritz, 2009; Tuch, Bargas-Avila, & Opwis, 2010). Utility and usability aspects such as an efficient task accomplishment have been emphasized. Nevertheless, the past decade has seen the emergence of a discussion of nonutilitarian aspects such as beauty, aesthetics, enjoyment, and fun (e.g., Blythe & Wright, 2003; Hassenzahl, 2004; Lavie & Tractinsky, 2004; Norman, 2004; Tractinsky, Shoval-Katz, & Ikar, 2000; van der Heijden, 2003). Under the umbrella term "user experience", the task-related view of system usability is extended through non-instrumental, affective and experiential aspects (Hassenzahl & Tractinsky, 2006; van Schaik & Ling, 2009). "Joy of use" as a concept of user experience explicitly addresses the role of affective design elements and the corresponding emotional reactions of the users (e.g., Hassenzahl, Burmester, & Beu, 2001; Reeps, 2006). It is assumed that objective system qualities such as layout, content, structure, or website design lead to a subjective perception of these qualities (e.g., appealing, usable), an emotional response in the user (e.g. pleasure, satisfaction) and to behavioral responses (e.g., approach, avoidance) (Hassenzahl, 2004). The focus of this line of research is on positive emotions.

1.1. Evoking positive emotions

But how can we design a website that can evoke positive emotions in learners? Even though the discussion surrounding emotional design factors has a longer tradition in web design than in multimedia learning, there is a lack of concrete, empirically validated design factors. However, *visual aesthetics* (as a more holistic design factor) has been shown to be a strong determinant of the subjective perception of a website and of the emotional responses for the users such as pleasure and enjoyment (e.g., Lavie & Tractinsky, 2004; Moshagen & Thielsch, 2010; Schenkmann & Jönsson, 2000; Tuch et al., 2010; van der Heijden, 2003). Lavie and Tractinsky (2004) identified two dimensions to the perception of website aesthetics: classical and expressive aesthetics. The classical aesthetics dimension refers to a clear, orderly alignment that appears clean, pleasant and symmetrical. It is therefore related to traditional notions of aesthetics as well as to design rules in the usability literature. The expressive aesthetics dimension refers to the novelty and unconventionality of a website. It depends on the designer's creativity and is associated with originality, fascination, and the use of special effects.

As visual aesthetics of websites are a rather holistic design feature, we still need to deduce more concrete design features that may evoke classical or expressive aesthetic impressions in users and according emotional responses. The visual aesthetics of websites has been shown to be affected by the *color* and *color combinations* used (e.g., Coursaris, Swierenga, & Watrall, 2008; Hartmann, Sutcliffe, & De Angeli, 2007; Thielsch, 2008). Moshagen and Thielsch (2010) identified color as one of four core dimensions of website visual aesthetics (alongside with simplicity, diversity and craftsmanship). Treating color as a concrete aesthetic design feature has the advantages that (a) color is an intrinsic design feature that does not require the incorporation of additional – potentially distracting – design elements (seductive details as mentioned above), (b) color has been shown to affect emotions and behavior in a wide variety of contexts (e.g., Elliot, Maier, Binser, Friedman, & Pekrun, 2009; Kaya & Epps, 2004; Valdez & Mehrabian, 1994; Weller & Livingston, 1988), (c) colors – in contrast to many other design features – can be completely characterized in terms of hue (chromatic tonality, e.g., blue, red), lightness and chroma (saturation) (cf. Valdez & Mehrabian, 1994), and (d) colors for webpages can be systematically varied by manipulating the proportion of red, green and blue in the RGB color model. Liedl (1994) suggests three techniques for achieving an aesthetically pleasing color harmony that is independent from individual preferences: (1) choosing complementary colors (opposing color pairs on the color wheel such as red/cyan, blue/yellow or green/magenta), (2) maximizing contrasts (by manipulating lightness and chroma, the complementary colors are made as unequal as possible – one very light/slightly saturated and the other one very dark/fully saturated), and (3) choosing similar colors (variations of one color by manipulating lightness and chroma). In a recent study, Müller, Heidig, & Reichelt (subm.) showed that systematic color designs based on these techniques (see Fig. 1 for examples) significantly differed in their perceived classical and expressive aesthetics as rated by the participants. They further affected the emotional states of the participants as their presentation resulted in significantly different ratings for valence and arousal, as measured using the "Self-Assessment Manikin" (SAM, Bradley & Lang, 1994).

1.2. Evoking negative emotions

Because this study aims to investigate both positive and negative emotions, we also need to deduce design features that may evoke negative emotional reactions. Naturally the web design literature does not explicitly contain considerations how to evoke negative emotions, instead providing suggestions on how to avoid them. Interestingly, the user experience literature dealing with affective issues of web design is almost exclusively focused on positive emotions. However, preventing negative emotional reactions such as dissatisfaction or frustration is a core aim of the usability literature (Hassenzahl & Tractinsky, 2006). Usability is defined as the extent to which a product such as a website can be used to reach a certain goal effectively, efficiently and satisfyingly (cf. ISO 9241–11). Negative emotions may occur when an inhibiting condition interferes with this goal (Lazar, Jones, Hackley, & Shneiderman, 2006). Inhibiting conditions and therefore causes of negative affects as named by users include interruptions due to bugs, system crashes, error messages and pop-ups, as well as hard-to-find features, auto formatting and long loading times (Ceaparu, Lazar, Bessiere, Robinson, & Shneiderman, 2002; Mentis & Gay, 2003). Out of these, long loading times seem to be the design feature most relevant to inducing negative emotions in multimedia learning, as they are (a) intrinsic in the sense that their manipulation does not require changes in the learning content itself, (b) easy to manipulate and (c) allow for a systematic variation. Previous studies showed that time delays lead to negative affect such as frustration and impatience, especially if no information about the length of the loading time is provided (Ceaparu et al., 2002; Dellaert & Kahn, 1999).

In the literature, suggested thresholds for intolerable loading times differ greatly and range from approximately 2–41 s (e.g. Bouch, Bhatti, & Kuchinsky, 2000; Nah, 2004; Ramsay, Barbesi, & Preece, 1998; Selvidge, Chaparro, & Bender, 2000). The most frequent citations is Nielsen's (1997) suggestion that users may lose interest or believe an error has occurred if loading times exceed 10 s. Galletta, Henry, McCoy, and Polak (2002) found decreases in behavioral intentions and performance when time delays exceeded 4 s, and decreases in attitudes for delays exceeding 8 s (cf. Nah,

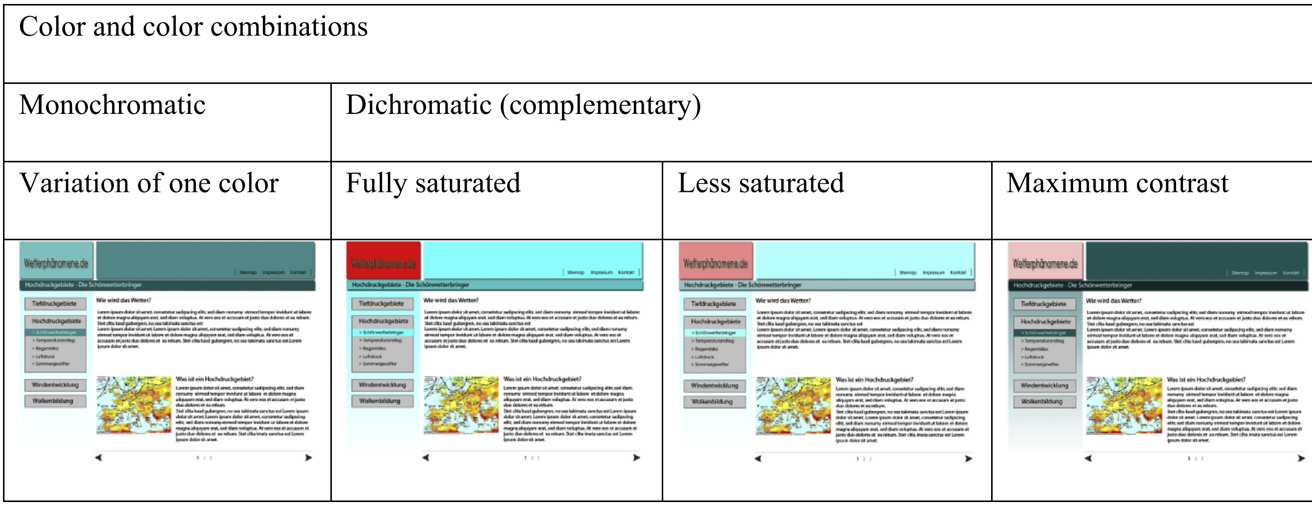

Fig. 1. Examples of systematically derived color combinations as used by Müller et al. (subm.). Variations of the color cyan and the color combination cyan/red applied to a multimedia learning material on weather phenomena.

2004). Bouch et al. (2000) asked their participants to press a button if they found a webpage loading time unacceptable. They report an average tolerance of 8.57 s. Selvidge et al. (2000) report significant differences in frustration levels between 1 and 30 s but no effects for 1, 10, and 20 s delays. However, the level of frustration experienced can be affected by many factors such as the user's expectations and experiences, the importance and nature of the task, the frequency of occurrence, the amount of time or work lost, and the related time pressure (Bessiere, Newhagen, Robinson, & Shneiderman, 2006; Ceaparu et al., 2002; Dellaert & Kahn, 1999; Lazar et al., 2006). Bouch et al. (2000) also identified the length of interaction time and method of page loading as relevant factors. They report that the user's tolerance for time delays decreased with increased time spent interacting with the system (cumulated slowness). Incremental loading was associated with more tolerance of latency compared to the display of the entire page.

1.3. Summary

Overall, the web design literature allows for the deduction of more holistic design features that may induce emotions in learners. Visual aesthetics (classical, expressive) seem to be appropriate for inducing positive emotions in learners. At a more concrete level, aesthetic color and color combinations can be used in order to evoke positive emotional reactions. For internal induction of negative emotions, usability can be applied as a holistic design feature, while the loading times for the learning material can be chosen as a concrete design feature. Both color and loading times can be considered as intrinsic design features, as they do not require any changes in the learning environment. Our research questions, however, not only aims to gain further insights into whether intrinsic design features of the multimedia learning material affect learners' emotions; the intent is also to investigate how emotions that are experienced during learning affect the learning process.

2. Emotions and learning

Previous research has demonstrated a relationship between learners' emotions and learning (e.g., Ainley, 2006; Goetz et al., 2012; Linnenbrink-Garcia, Rogat, & Koskey, 2011). For instance, Pekrun, Goetz, Titz, and Perry (2002) found that pleasant emotions such as enjoyment predict high achievement while unpleasant emotions such as test anxiety predict low achievement. Pekrun, Götz, Frenzel, Barchfeld, & Perry (2011) showed that learners' achievement emotions were linked to their control and value appraisals, motivation, self-regulation, and academic performance. Reviewing the literature, terms such as affect, mood, and emotion are often used synonymously (Pekrun, Elliot, & Maier, 2009). To describe these psychological constructs, many authors distinguish between mood and emotion based on intensity and duration. Moods are longer lasting and more diffuse, without a particular referent (Pekrun, 2006; Rosenberg, 1998). In contrast, emotions consist of short, intense episodes in response to a specific referent (Rosenberg, 1998; Schwarz, 1990; Schwarz & Clore, 1996). Affect is used as a broader term covering mood and emotion, and is defined as the subjective experience of emotional episodes (Bagozzi, Gopinath, & Nyer, 1999; Frijda, 1999; Watson, Clark, & Tellegen, 1988). Negative affect describes the extent of negative and unpleasant emotions while positive affect describes the extent of positive and pleasant emotions (Um et al., 2012; Zhang & Lee, 2009). In our study, we also refer to "affect" as an undifferentiated emotional state (e.g., Fredrickson, 2003; Pintrich, 2000; Tellegen, Watson, & Clark, 1999). Furthermore, we describe emotions along two dimensions that affect performance: valence (positive–negative) and activation (activating–deactivating). The assumption of a two-dimensional differentiation is widely used in the literature (e.g., Russell, 2003). However, previous studies into emotional design in multimedia learning (e.g., Plass et al., 2014; Um et al., 2012) have applied the PANAS scale, which includes positive and negative activation (Tellegen et al., 1999; Watson & Tellegen, 1985). To choose a more differentiated view on learners' emotions, we used the PANAVA scale (Schallberger, 2005), which assesses valence as additional dimension.

2.1. Considering emotions in multimedia learning

Although a range of research (e.g., Artino, 2012; Chanlin, 1998; Linnenbrink, 2006; Pekrun et al., 2009) indicates that affective state (i.e., considering a learner's emotional state) is an important factor in designing learning materials, only few empirical studies exist focusing on learner emotions and how they affect learning performance in multimedia learning (e.g., Plass et al., 2014; Um et al., 2012). In examining how emotions experienced during learning may affect the learning process, two contrary lines of evidence can be differentiated: on the one hand, it is assumed that positive emotions can require additional task processing or task-irrelevant processing, and negatively influence learning (suppression hypothesis: Oaksford, Morris, Grainger, & Williams, 1996; Seibert & Ellis, 1991). This assumption can be explained within *cognitive load theory* (CLT, e.g., Paas, Renkl, & Sweller, 2003). Since the resources of working memory are limited, emotions experienced during cognitive processing of learning materials impose a needless load on working memory (Chen & Wang, 2011; Seibert & Ellis, 1991). On the other hand, positive emotions have been examined as factors facilitating changes in affective components such as motivation, creativity, and problem-solving skills (Chen & Wang, 2011; Um et al., 2012). Under this line of argument, positive emotions enhance long-term memory and retrieval, and facilitate working memory processes (facilitation hypothesis, Erez & Isen, 2002; Isen, 2000; Norman, 2002). Even though the facilitation hypothesis is prevalent in research into positive emotions, the influence of learners' positive emotional state on complex learning is not yet clear. In the context of multimedia learning, Um et al. (2012) showed that applying emotional design principles to multimedia learning materials can induce positive emotions. Further, positive emotions facilitated cognitive processes and learning. Each of the 34 college students who participated in this study was given one of four study designs that were manipulated by two factors: externally induced positive or neutral emotions along with a positive or neutral design of the learning material. The neutral and positive multimedia material both explained how immunization works using the same learning material over the equal length of time, but they were presented in different design aesthetics. The neutral one had colorless material (gray-scale) with no character design and the positive one was colorful with anthropomorphic characters (molecules in the learning unit were decorated with human faces). The results showed that a multimedia learning environment with positive emotional design (warm colors and face-like shapes) was able to evoke positive emotions in learners and resulted in higher comprehension and transfer performance compared to the neutral design (gray colors and no face-like shapes). Further, Um et al. (2012) found that learners using the positive emotional design perceived the learning materials as less difficult, invested more mental effort, and reported higher levels of motivation and satisfaction. Using the same learning material, Plass et al. (2014) also examined whether design factors such as color and shape in multimedia learning material induce positive emotions in learners, and how they affect cognitive outcomes (e.g., cognitive load, learning outcomes) as well as affective outcomes (e.g., motivation, satisfaction). The results of study 1 showed that well-designed materials evoke positive emotions and facilitate comprehension, reduce the perceived task difficulty, and increase the levels of learners' motivation. However, transfer was not affected by emotional design. Study 2 found that round face-like shapes both alone and in combination with warm colors induced positive emotions. However, warm colors alone, did not influence learners' emotions. Comprehension was facilitated through warm colors and round face-like shapes, whereas transfer was facilitated using neutral colors and round face-like shapes. These experiments provide a different and novel view on emotional design in multimedia learning compared to the seductive details effect (Harp & Mayer, 1997; Mayer et al., 2001), which assumes that appealing – but irrelevant – learning materials harm learning (Mayer, 2009). The learning environments in the presented studies (Plass et al., 2014; Um et al., 2012) were aesthetically designed to evoke positive emotions. Since the influence of negative emotional states has rarely been considered (e.g., Goetz, Pekrun, Hall, & Haag, 2006) and not yet been considered in relation to an emotional design for multimedia learning, in our study we also examined the impact of negative emotional states on the learning process. To do so, the manner in which emotions (positive vs. negative) can affect complex learning processes must be clarified. Empirical findings indicate that positive emotional states promote creative, flexible, and intuitive-holistic ways of thinking (top-down) (Bless, 2001; Bless & Igou, 2006), whereas negative emotional states promote analytical–sequential, detailed, and rigid ways of processing information (bottom-up) (Ashby, Isen, & Turken, 1999; Fiedler, 1988; Pekrun et al., 2009). Both the breadth and depth of processing are affected (Bless, 2001; Bless & Igou, 2006). Hence, positive emotional states may facilitate comprehension and transfer performance, whereas negative emotional states may facilitate recall performance.

2.2. Considering the role of motivation

Emotional states not only affect information processing, but can also induce and modulate learners' interest and motivation to learn (Erez & Isen, 2002; Linnenbrink, 2006; Pekrun, 2006). For instance, Meyer and Turner (2007) posited that emotions might change the way that learners perceive the educational experience and thus change their ensuing motivation. Moreover, they hypothesize that the affective state of interest might mediate the relationship of achievement goal orientations, engagement to learn, and efficacy in cognitive processing. Particularly, emotions influence how learners engage in the activity and their motivation to persist in the face of difficulties. According to Park and Lim (2007), illustrative types of multimedia learning materials effect learners' motivation. Further, they found that learners pay more attention to learning materials when emotionally interesting illustrations are presented than to a text-only information condition. Additionally, learners in the emotional interest condition showed significant higher relevance scores compared to the text-only group. For complex learning processes, the cognitive-motivational mediation model by Pekrun (1992) assumes that positive emotions promote intrinsic motivation. Intrinsic motivation can be defined as 'the motivation to engage in a task for its own sake – out of interest and/or enjoyment – and not as a means to another reward' (Isen & Reeve, 2005, p. 298). Experiments by Isen and Reeve (2005) confirmed that positive emotions foster intrinsic motivation, as shown by choice of activity in a free-choice situation and by rated level of enjoyment of a novel and challenging task. In reference to the research field of emotional design, these findings are also confirmed by Um et al. (2012): learners using a positive emotional design reported higher motivation than learners receiving a neutral design.

In summary, learners' emotional states are not generally expected to promote the learning process; instead, the impact of positive and negative emotional states should be investigated based on the requirements and learning goals (analytical vs. holistic). To specify our research hypotheses, these differentiations were considered.

3. Current research gaps and aims of this study

Current research into designing multimedia learning material often focuses on learning outcomes and cognitive aspects (e.g., cognitive load). A weak point of previous studies is that emotional aspects were not or only rarely considered. First, the presented studies seek to provide more insights into an emotional design of multimedia learning by systematically investigating design features that may evoke emotional states in learners. Taking a multidisciplinary approach by including research results from the field of web design, we were able to systematically deduce design features that may evoke both positive and negative emotional states. Second, in contrast to existing studies on emotional design in multimedia learning, this study examines emotional aspects in more detail: we are not only looking at positive emotional states but also investigating the effect of negative emotional states. In reference to the literature of fundamental research, we were able to derive more differentiated hypotheses for the relation between the learners' emotional states and the learning outcomes. We have also chosen a more differentiated view of the assessment of learners' emotional states. Whereas the PANAS scale used in the previous studies of emotional design (Plass et al., 2014; Um et al., 2012) assesses positive and negative activation (Tellegen et al., 1999), we use the PANAVA scale (Schallberger, 2005), which also assesses valence as a dimension of learners' emotional experience. Third, previous studies on emotional design used both an external emotion induction (via the Velten method or via films) and an internal emotion induction (via the design of the learning material). However, it is hard to distinguish the effects of the initial external emotion induction and the internal emotion induction while using the learning material throughout the learning process. Therefore, we only used internal emotion induction via the design of the learning material. This also adds to the external validity of the studies.

4. Research questions and hypotheses

Our study aims to systematically analyze the impact of different design factors on learners' emotional states and the role of emotional states in complex learning processes. To fill the aforementioned research gaps, this study examined the following hypotheses.

4.1. Hypothesis 1

Do intrinsic design features of the multimedia learning material affect learners' emotional states? Previous research (Plass et al., 2014; Um et al., 2012) found that an appealing design evoked more positive emotional states than a neutral one. Our study sought to replicate these empirical findings with different intrinsic design features, a more complex learning material and other learning content in order to gain further insights into the stability of the effect. We systematically deduced intrinsic design elements that may evoke positive and negative emotional states in learners. Different color schemes were used in order to manipulate the aesthetics of the multimedia learning material. The resulting two design factors (classical vs. expressive aesthetics) with two levels each (high vs. low) were expected to evoke more positive emotional states. In the hope of encouraging negative emotional states in learners, we manipulated the page load times of the learning material as a usability factor (high vs. low usability), and compared these experimental conditions to a control group (gray scale). Based on the literature review, we hypothesized that highly aesthetical designs (classical, expressive) would result in more positive emotional states than low aesthetical designs (hypothesis 1.1). Further, we assumed that low usability (longer loading times) would induce more negative emotional states than high usability (hypothesis 1.2).

4.2. Hypothesis 2

How do emotional states experienced during learning affect the learning process? A literature review shows that this question is controversial in current research. We agree with existing studies (e.g., Bless & Igou, 2006; Pekrun et al., 2009) which claim that positive emotional states can facilitate an intuitive, holistic information processing (top–down), while negative emotional states can facilitate an analytical and sequential processing (bottom–up). Previous research (Plass et al., 2014; Um et al., 2012) found that positive emotional states can facilitate cognitive processes and learning. In reference to complex learning processes, positive emotional states may enhance comprehension and transfer performance, along with learners' intrinsic motivation (e.g., Isen 1984; Isen, Johnson, Mertz, & Robinson, 1985). With regard to the different measures of learning outcomes (retention, comprehension, and transfer), we provide differentiated hypotheses: positive emotional states in learners facilitate comprehension and transfer performance (hypothesis 2.1), while negative emotional states foster retention (hypothesis 2.2). Based on the literature review reported above, we assume positive emotional states facilitate intrinsic motivation (hypothesis 2.3), while negative emotional states hamper intrinsic motivation (hypothesis 2.4).

5. Method

5.1. Participants and design

The participants were 334 college students (273 women, 61 men) majoring in education, psychology or other social studies at a German university. The mean age was 22.39 years (SD = 2.10). Participants were randomly assigned to one of nine conditions, created by two design factors (classical vs. expressive aesthetics) with two levels each (high vs. low) and one usability factor (high vs. low usability) as well as a control group (no color/gray scale). This resulted in the following nine conditions: classical aesthetics high/usability high (CHUH, $n = 38$), classical aesthetics high/usability low (CHUL, $n = 37$), classical aesthetics low/usability high (CLUH, $n = 36$), classical aesthetics low/usability low (CLUL, $n = 35$), expressive aesthetics high/usability high (EHUH, $n = 39$), expressive aesthetics high/usability low (EHUL, $n = 35$), expressive aesthetics low/usability high (ELUH, $n = 38$), expressive aesthetics low/usability low (ELUL, $n = 40$), and control group ($n = 36$). The experimental design is shown in Table 1.

5.2. Pre-study: designing the learning material

A pre-study was conducted in order to identify designs for the learning material that could be used as high vs. low in either classical aesthetics (clear, clean, pleasant) or expressive aesthetics (original, fascinating, creative). The participants received a computer-based multimedia learning environment explaining how weather works. The computer program was developed using Adobe Captivate software. The learning content was presented in four chapters. The 60 pages of the learning material incorporated text, pictures, and short animations. The material was self-paced, allowing the learner to choose when to continue to the next page but not change the order of the presentation of the pages.

Within the pre-study, 173 college students rated 32 designs for the learning material on classical and expressive aesthetics using the "Measurement Instrument of Perceived Visual Aesthetics of Websites" (MIPVA, Lavie & Tractinsky, 2004). The 32 designs for the learning material differed in their color design, which was varied using three methods to achieve color harmony (Itten, 1989; Liedl, 1994): complementary colors, similarity of colors and the

Table 1
Experimental design of the study.

	Usability		Control
	High	Low	
Classical aesthetics			
High	CHUH, $n = 38$	CHUL, $n = 37$	CONTROL, $n = 36$
Low	CLUH, $n = 36$	CLUL, $n = 35$	
Expressive aesthetics			
High	EHUH, $n = 39$	EHUL, $n = 35$	
Low	ELUH, $n = 38$	ELUL, $n = 40$	

method of maximum contrasts. Further color gradients and shining effects were used in order to create designs high in expressive aesthetics (cf. Müller et al., subm.). The designs that were rated highest and lowest on either classical or expressive aesthetics were chosen as stimuli for the main study (see Fig. 2). A yellow design (similarity of colors) served as the high classical aesthetics condition ($M = 6.29$, SD = 1.38 on the classical aesthetics scale ranging from 1-low to 9-high). A magenta-green design (complementary colors) served as low classical aesthetics condition ($M = 3.14$, SD = 1.37). The ratings of these two designs differed significantly along the classical aesthetics dimension ($F(31, 1006) = 6.01$, $p = .006$, $\eta^2 = .16$). Next, a yellow–green color gradient design served as the high expressive aesthetics condition ($M = 6.22$, SD = 1.57) and a red–cyan design (complementary colors) served

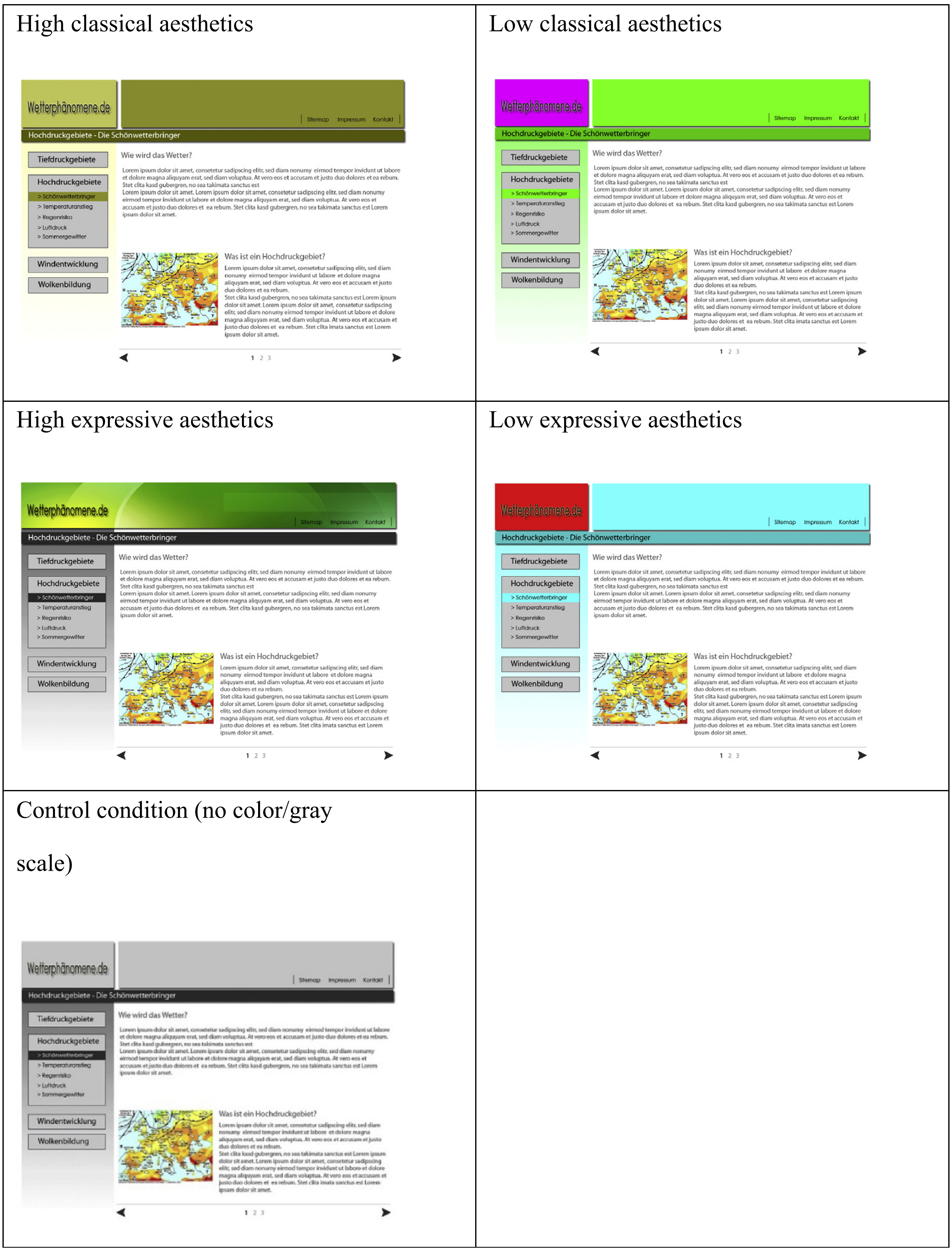

Fig. 2. Color designs for the multimedia learning material on how weather works. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

as the low expressive aesthetics condition ($M = 3.75$, SD = 1.68). The ratings of the two expressive aesthetics designs differed significantly along the expressive aesthetics dimension ($F(31, 1006) = 3.03$, $p < .001$, $\eta^2 = .09$).

In order to create a high and low usability condition, we manipulated the loading times of the pages of the learning material. In line with current suggestions on thresholds for inacceptable loading times (e.g., Bouch et al., 2000; Galletta et al., 2002), we chose an 8.5 s limit. Due to the length of the interaction (60 pages of learning material to load) we decided on a limit below 10 s. Hence, each page in the low usability condition needed an extra 8.5 s to load compared to the high usability condition. Furthermore, the participants did not receive any information on how long they would have to wait, and the pages did not load incrementally, instead appearing in their entirety after the loading time.

5.3. Measures

In order to assess the *perceived aesthetics* of the learning material as a control variable, the "Measurement Instrument of Perceived Visual Aesthetics of Websites" (MIPVA, Lavie & Tractinsky, 2004) was applied. It consists of the scales "classical aesthetics" (Cronbach's $\alpha = .81$) and "expressive aesthetics" (Cronbach's $\alpha = .86$). Each scale contained five items that were rated on a 6-point Likert-type scale ranging from 1 (not at all) to 6 (very much). In completing the MIPVA, participants indicated the degree to which they perceived the design of the learning material as "clear", "clean", "aesthetic", "symmetrical", or "pleasant" (classical aesthetics) and "original", "fascinating", "sophisticated", "creative", or "uses effects" (expressive aesthetics).

Perceived usability was employed as a second control variable, and was assessed using the usability scale from of the "Questionnaire User Experience (QUX)" (Müller, 2009; Müller, Heidig, & Niegemann, 2012). The 7 items (e.g., "I am able to quickly reach my goals.") were rated on a 6-point Likert-type scale ranging from 1 (not at all) to 6 (very much) (Cronbach's $\alpha = .87$).

The learners' *emotional states* were measured using the "Scales for Assessing Positive/Negative Activation and Valence in Experience Sampling Studies" (PANAVA-KS; Schallberger, 2005). The PANAVA-KS is based on the "Positive and Negative Affect Schedule" (PANAS; Watson et al., 1988), an established measure of current affect that was also used in prior studies into emotional design in multimedia learning (cf. Plass et al., 2014; Um et al., 2012). The PANAVA-KS was preferred over the PANAS for four reasons: (1) conceptual clarity, (2) inclusion of the "valence" dimension, (3) use of bipolar items and (4) brevity. (1) Conceptual clarity: the positive affect scale of the PANAS does not solely comprise positive states but positive states with high activation and negative states with low activation. Analogously, the negative affect scale comprises negative states with high activation and positive states with low activation. Tellegen et al. (1999) therefore renamed them as "positive and negative activation", where "positive activation" represents "goal-directed behaviors" and "negative activation" represents "withdrawal behaviors" (Tellegen et al., 1999). Although this view is shared by different authors in the current literature (e.g., Cacioppo, Gardner, & Berntson, 1999; Gable, Reis, & Elliot, 2003; cf. Schallberger, 2005), the PANAS is still frequently applied and interpreted as a measure of positive and negative affect. (2) Inclusion of the valence dimension: as well as the "positive activation" and "negative activation" scales measured by the PANAS, the PANAVA-KS additionally contains valence (pleasantness, Russell, 1980) as a third dimension. (3) Use of bipolar items: in order to avoid extreme skewness of the distribution, the PANAVA-KA contains bipolar items such as unhappy–happy (valence), bored–excited (positive activation), quiet–angry (negative activation). (4) Brevity: the PANAS incorporates 10 items each for positive and negative activation, whereas the PANAVA-KS includes only 2 items for valence and 4 items each for positive and negative activation. It is therefore more suited to repeated measurement designs such as those used in the presented study. The 10 items of the PANAVA-KS were rated on a sliding bar representing a 100 point scale ranging from 1 (low) through 50 (neutral) to 100 (high). Verbal anchors were only provided for the extremes and the neutral position (e.g., "unhappy" very much – neutral – very much "happy").

In order to assess *prior knowledge*, the learners had to answer 6 multiple-choice questions (e.g., "What does the term 'climate' mean?", "What is the troposphere?"). The total score for each participant was obtained by adding points for all questions, resulting in a range from 0 to 13 points. The questions presented for the prior knowledge test were not part of the post-test.

The *learning outcomes* test consisted of retention, comprehension, and transfer tasks. The *retention* test measured the degree to which the learners were able to recall and reproduce facts that were presented in the material. It consisted of 12 questions, including 6 multiple-choice questions (e.g., "What does the term 'weather' mean?"), 5 requests to name key concepts or their elements (e.g., "Please name three parameters that affect our weather!"), and 1 cloze ("Several weather services have been developed since ___, ..."). The participants received 1 point for each correct answer, resulting in a range of 0–19 points. The *comprehension* test measured learners' understanding of key concepts within the material. It contained seven questions, including 5 multiple-choice questions (e.g., "Why is the weather in anticyclones mainly nice?") and two requests to explain weather phenomena and processes (e.g., "Why doesn't it rain on the moon?"). The participants received 1 point for each correct answer, resulting in a range of 0–13 points. The *transfer* test measured learners' ability to apply the learned concepts to solve problems. It consisted of seven questions, including 6 multiple-choice questions (e.g., "Please indicate which area (areas A and B, or both) in the isobaric chart (a) is affected by a cold front, (b) has more wind, (c) has more clouds ...") and one open question ("Which of the two towns in the figure has higher temperatures? Please give reasons for your answer!"). The participants received 1 point for each correct answer, resulting in a range of 0–15 points.

The learners' *intrinsic motivation* was assessed using the "self-report measure of intrinsic motivation" (Isen & Reeve, 2005). It incorporates 8 items (e.g., "I would like to continue working with the learning material.") that were rated on a 6-point Likert-type scale ranging from 1 (not at all) to 6 (very much) (Cronbach's $\alpha = .93$).

5.4. Procedure

Participants were tested in groups of 12 in a computer lab, where each participant worked individually on one computer. The procedure was entirely administered through a computer system: the questionnaires were directly included in the program, and the participants answered it online.

After receiving an introduction to the computer-based procedures, participants were asked to follow the instructions on the screen. The session then started with the demographic questionnaires, followed by the first presentation of the PANAVA-KS in order to assess the baseline of the participants' emotions as a control variable (baseline PANAVA). Next, the learning material was presented to the participants in one of the nine designs, which varied in their aesthetic design and usability. Participants were told that they would have to fill out a learning test after studying the material. In order to assess the learners' current emotional state, the PANAVA-KS was presented twice, first in the middle of the learning material (after 30 pages) and again at the end (after 60 pages) (PANAVA 2 and 3). This was followed by the questionnaires on perceived aesthetics, perceived usability, intrinsic motivation, and the learning outcomes test (retention, comprehension, and transfer tasks). Learning time was assessed via log-files. This was of particular importance so as to account for differences in time spent on the material due to the loading times generated by the usability conditions. Overall, the experiment lasted about 90 min. Participants took on average 31.04 min (SD = 6.54) to study the 60 pages of the learning material, depending on their working speed (self-paced instruction).

6. Results

Table 2 shows the means and standard deviation for the perceived classical and expressive aesthetics and usability in each of the experimental conditions.

6.1. Manipulation check

6.1.1. Manipulation check for the classical aesthetics conditions

We first investigated whether the manipulation of the classical aesthetics has been successful. To do this, we considered the high and low classical aesthetics groups as along with the control group (CHUH, CHUL, CLUH, CLUL, control group). A one-way ANOVA was conducted with the classical aesthetics conditions (high vs. low vs. control group) as between-subjects factor, and the perceived classical aesthetics as rated by the participants as dependent measure. The analysis revealed no significant differences between the experimental conditions ($F(4, 177) = 1.26$, MSE = 0.91, $p = 0.286$, $\eta^2 = 0.028$). This result suggests that the manipulation of the classical aesthetics was unsuccessful, as the participants did not perceive the different designs of the multimedia learning material as different in their appeal along the classical aesthetics dimension.

6.1.2. Manipulation check for the expressive aesthetics conditions

Next, we analyzed whether the manipulation of the expressive aesthetics has been successful. To do this, we investigated the high and low expressive aesthetics groups as well as the control group (EHUH, EHUL, ELUH, ELUL, control group). We conducted a one-way ANOVA with the expressive aesthetics conditions (high vs. low vs. control group) as between-subjects factor, and the perceived expressive aesthetics as rated by the participants as dependent measure. It yielded no significant differences between the experimental conditions ($F(4, 183) = .20$, MSE = .20, $p = .941$, $\eta^2 = .004$). This result indicates that the manipulation of the expressive aesthetics was also unsuccessful.

6.1.3. Manipulation check for the usability conditions

Third, we examined whether manipulating usability by controlling the loading times of the pages of the learning material had been successful. To do this, we compared the perceived usability as rated by the participants in the high usability groups (CHUH, CLUH, EHUH, ELUH) to the low usability groups (CHUL, CLUL, EHUL, ELUL). A one-way ANOVA was conducted with the usability conditions (high vs. low) as between-subjects factor, and the perceived usability as rated by the participants as dependent measure. The analysis again revealed no significant differences between the experimental conditions ($F(1, 296) = .22$, MSE = .08, $p = .636$, $\eta^2 = .001$). Based on this result, the manipulation of the usability of the learning material was also unsuccessful.

6.1.4. Interim discussion

Contrary to our expectations, the manipulation checks of the two aesthetics conditions and the usability condition revealed that our manipulations were unsuccessful. This was an unexpected result as a thorough pre-study was conducted in order to find color designs that would be perceived as high vs. low in either classical or expressive aesthetics. This pre-study compared 32 designs that were tested with a large sample size ($N = 173$). The result for the usability manipulation was even more surprising. In the low usability condition, each of the 60 pages of the learning material needed an extra 8.5 s to load compared to the high usability condition. However, the perceived usability of the learning material did not differ between those two experimental conditions. While a detailed discussion of these results will be given in the general discussion section, the next steps of our analyses need to be discussed in light of the failed manipulation checks. Without successful manipulation checks, we lack the basis for the between-group comparisons that would have been the first intended method of analysis. Our manipulation checks were based on perceived aesthetics and usability. One could therefore argue that the differences – at least for the usability manipulation – are nevertheless objectively given. We therefore decided to report the results of the between-group comparisons in order to thoroughly present and discuss the data. However, the results need to be interpreted with the failed manipulation checks in mind.

6.2. Results of the between-group comparisons

6.2.1. Emotional states of the learners

Table 3 shows the means and standard deviation for the learners' emotional states at the three measuring times: before (baseline, PANAVA 1), during (PANAVA 2), and after the learning material (PANAVA 3) in each of the experimental conditions.

We first controlled for differences in the *learners' baseline emotional states* among the nine conditions. One-way ANOVAs with the learner's emotional state (valence, positive activation, and negative activation) as dependent measure and condition as factor revealed no significant differences (valence: $F(8, 325) = 1.07$, MSE = 290.38, $p = .381$, $\eta^2 = .026$; positive activation: $F(8, 325) = 1.08$, MSE = 175.43, $p = .378$, $\eta^2 = .026$; negative activation: $F(8, 325) = 1.08$, MSE = 271.35, $p = .375$, $\eta^2 = .026$). These results

Table 2
Means and standard deviations for each experimental condition on perceived classical and expressive aesthetics as well as perceived usability.

Classical aesthetics				Expressive aesthetics				Control
CHUH M (SD)	CHUL M (SD)	CLUH M (SD)	CLUL M (SD)	EHUH M (SD)	EHUL M (SD)	ELUH M (SD)	ELUL M (SD)	M (SD)
Perceived classical aesthetics								
4.61 (.72)	4.36 (1.04)	4.18 (.86)	4.44 (.72)	4.57 (.74)	4.55 (.77)	4.44 (.85)	4.46 (.94)	4.47 (.86)
Perceived expressive aesthetics								
4.15 (.84)	3.85 (1.06)	4.08 (.95)	4.39 (.90)	4.27 (1.11)	4.10 (.94)	4.30 (1.02)	4.23 (1.02)	4.22 (.98)
Perceived usability								
5.52 (.58)	5.43 (.55)	5.38 (.73)	5.57 (.44)	5.61 (.46)	5.52 (.43)	5.45 (.84)	5.57 (.49)	5.41 (.89)

Note: Potential scores ranged from 1 (not at all) to 6 (very much).

Table 3
Means and standard deviations for each experimental condition on the learners' emotional states before (baseline, PANAVA 1), during (PANAVA 2), and after the learning material (PANAVA 3).

	Classical aesthetics				Expressive aesthetics				Control
	CHUH M (SD)	CHUL M (SD)	CLUH M (SD)	CLUL M (SD)	EHUH M (SD)	EHUL M (SD)	ELUH M (SD)	ELUL M (SD)	M (SD)
Valence 1	62.68 (15.82)	58.49 (11.63)	60.71 (16.83)	63.87 (17.65)	57.42 (16.93)	63.49 (14.91)	64.30 (18.28)	61.80 (16.49)	57.14 (18.36)
Positive activation 1	51.05 (11.80)	49.66 (13.02)	52.91 (11.92)	54.99 (14.94)	53.38 (11.39)	56.33 (11.44)	52.68 (12.47)	52.84 (11.10)	49.83 (16.16)
Negative activation 1	40.27 (16.25)	41.46 (15.83)	39.22 (17.30)	36.68 (13.79)	40.37 (17.60)	36.71 (15.18)	38.74 (15.48)	45.04 (16.34)	42.56 (14.00)
Valence 2	58.29 (13.68)	56.33 (10.56)	57.32 (16.68)	64.56 (13.10)	58.73 (12.44)	60.81 (11.23)	64.03 (15.00)	60.50 (14.40)	57.69 (12.22)
Positive activation 2	48.80 (14.08)	45.07 (14.37)	49.53 (16.65)	54.94 (16.29)	52.65 (14.09)	55.08 (13.07)	53.36 (12.07)	51.43 (11.82)	47.50 (14.31)
Negative activation 2	34.64 (16.64)	35.53 (14.20)	34.84 (14.06)	32.21 (14.84)	30.88 (14.45)	35.06 (14.30)	32.33 (13.94)	38.65 (16.79)	34.52 (14.36)
Valence 3	60.26 (13.00)	56.73 (12.07)	59.19 (15.32)	65.49 (12.20)	60.53 (12.56)	62.10 (12.19)	62.84 (14.20)	60.46 (12.60)	60.58 (9.81)
Positive activation 3	47.19 (15.79)	46.30 (14.09)	50.66 (14.19)	55.87 (15.46)	51.42 (15.65)	53.22 (13.46)	54.57 (11.91)	51.31 (11.63)	48.67 (13.71)
Negative activation 3	33.86 (16.74)	34.58 (15.04)	34.38 (13.89)	30.38 (15.67)	31.68 (13.28)	30.54 (15.81)	31.55 (15.57)	36.16 (15.14)	36.72 (15.01)

Note: Potential scores ranged from 1 (low) across 50 (neutral) to 100 (high).

suggest that the emotional states prior to starting the learning material did not differ across the experimental conditions.

In order to investigate whether the nine different designs for the multimedia learning material were able to affect learners' emotional states we calculated ANOVAs with repeated measures (RM_ANOVAs) with the PANAVA scores at the baseline, in the middle of the learning material and after the learning material as repeated measures variable and the condition as between-subjects factor. The results indicate significant changes in the PANAVA scores over time (valence: Wilks' $\Lambda = 0.98$, $F(2, 324) = 3.33$, $p = .037$, $\eta^2 = .020$; positive activation: Wilks' $\Lambda = 0.98$, $F(2, 324) = 3.01$, $p = .051$, $\eta^2 = .018$; negative activation: Wilks' $\Lambda = 0.81$, $F(2, 324) = 37.08$, $p < .001$, $\eta^2 = .186$), suggesting that the valence scores dropped between the baseline and the middle of the learning material but then increased by the end of the learning material with a small effect size, whereas the positive and negative activation scores decreased from the baseline to the middle of the learning material and did not increase towards the end. The changes in positive activation were only marginally significant; however, the changes in negative activation yielded a medium to large effect size. The analysis further revealed no interaction between the changes in PANAVA scores and group (valence: Wilks' $\Lambda = 0.97$, $F(16, 648) = .72$, $p = .777$, $\eta^2 = .017$; positive activation: Wilks' $\Lambda = 0.96$, $F(16, 648) = .85$, $p = .629$, $\eta^2 = .021$; negative activation: Wilks' $\Lambda = 0.95$, $F(16, 648) = 1.05$, $p = .403$, $\eta^2 = .025$), suggesting that the changes in the PANAVA scores over time did not differ between the experimental conditions. Finally, there were no main effects of group on valence ($F(8, 325) = 1.56$, MSE = 682.94, $p = .136$, $\eta^2 = .037$) and negative activation ($F(8, 325) = 1.02$, MSE = 550.83, $p = .417$, $\eta^2 = .025$), indicating that the participants in the nine experimental conditions did not report different emotional states on these two dimensions. However, the analyses did yield a main effect of group for positive activation with a small effect size ($F(8, 325) = 2.15$, MSE = 889.87, $p = .031$, $\eta^2 = .050$). The Student–Newman–Keuls post hoc test indicated that the high and low classical aesthetics groups differed in their positive activation scores, whereas the other experimental groups did not show significant differences. We therefore calculated a follow-up two-way RM_ANOVA with classical aesthetics and usability as between-subjects factors and the PANAVA scores at the baseline, in the middle of the learning material and after the learning material as repeated measures. It revealed no main effect for usability ($F(1, 142) = 0.30$, MSE = 135.99, $p = .585$, $\eta^2 = .002$) and no interaction between classical aesthetics and usability ($F(1, 142) = 2.34$, MSE = 1061.26, $p = .128$, $\eta^2 = .016$) but a main effect for classical aesthetics ($F(1, 142) = 6.37$, MSE = 2888.42, $p = .013$, $\eta^2 = .043$), suggesting that learners in the low classical aesthetics groups reported higher positive activation than learners in the high classical aesthetics groups.

6.2.2. Interim discussion

These results indicate that different designs for the multimedia learning material did not affect the learners' emotional states in the expected way. The learners in the high aesthetic groups (classical, expressive) did not show more positive emotional states than the learners in the low aesthetics groups (hypothesis 1.1). The opposite was even shown for classical aesthetics. Furthermore, learners in the low usability groups did not report more negative emotional states than those in the high usability groups (hypothesis 1.2). Therefore, the objectively provided differences in usability and aesthetics did not affect the learners' emotional states in a systematic way. Due to the failed manipulation checks, these results cannot be used to interpret the perceived aesthetics or the perceived usability. Since there are no effects on learners' emotional states, the between-group comparisons cannot be used to answer the second research question. Thus the between-group comparisons of learning outcomes and intrinsic motivation are omitted.

6.3. Results of the regression analyses

Due to the failed manipulation checks, the between-group comparisons as described above only deal with objective differences in aesthetics and usability. We therefore applied regression analyses in order to investigate the effects of the perceived aesthetics and usability.

6.3.1. Research question 1: regression of the design of the multimedia learning material on the learners' emotional states

Regression analyses were conducted in order to test the hypothesis of whether the perceived design of the multimedia learning material affects participants' ratings of their emotional states during learning. We calculated single linear regression analyses with the perceived design of the learning material (perceived classical and expressive aesthetics and perceived usability) as independent variable and the learners' emotional states (valence, positive activation, negative activation) as dependent variables. As indicators for the learners' emotional states during learning, we used the mean of the ratings on the emotional states as assessed in the middle of the learning material and after the learning material (mean of PANAVA 2 and 3).

The results indicated that the participants' ratings of perceived classical aesthetics significantly predicted the ratings for their emotional states (valence: $\beta = .235$, $p<.001$, positive activation: $\beta = .314$, $p < .001$, negative activation: $\beta = -.175$, $p = .001$). The perceived classical aesthetics as rated by the learners explained 5.5% of the variance in valence ratings (R^2 = .055, F(1, 332) = 19.43, p < .001), 9.8% of the variance in positive activation ratings (R^2 = .098, F(1, 332) = 36.20, p < .001) and 3.1% of the variance in negative activation ratings (R^2 = .031, F(1, 332) = 10.51, p = .001). Participants who reported higher scores on perceived classical aesthetics also reported more positive emotional states: higher valence levels and positive activation and lower negative activation.

The results further showed that the participants' ratings for perceived expressive aesthetics significantly predicted the ratings for valence (β = .183, p = .001) and positive activation (β = .365, p < .001), but not for negative activation (F(1, 332) = .61, p = .435). The ratings for the perceived expressive aesthetics explained 3.3% of the variance in valence (R^2 = .033, F(1, 332) = 11.44, p = .001), and 13.4% of the variance in positive activation (R^2 = .134, F(1, 332) = 51.16, p < .001). Learners who perceived expressive aesthetics as higher reported marginally higher levels of valence and considerably higher levels of positive activation. However, the learners' negative activation remained generally unaffected by the perceived expressive aesthetics.

Finally, the results of the analyses suggested that the learners' ratings for perceived usability of the learning material also significantly predicted the ratings on their emotional states (valence: β = .148, p = .007, positive activation: β = .201, p < .001, negative activation: β = −.148, p = .007). The participants' ratings for perceived usability explained 2.2% of the variance in valence ratings (R^2 = .022, F(1, 332) = 7.44, p = .007), 4.1% of the variance in positive activation ratings (R^2 = .041, F(1, 332) = 14.05, p < .001) and 2.2% of the variance in negative activation ratings (R^2 = .022, F(1, 332) = 7.46, p = .007). Hence, participants who reported higher scores for perceived usability also reported more positive emotional states: higher levels of valence and positive activation and lower negative activation.

6.3.2. Interim discussion

In summary, these results support the hypothesis that the design of multimedia learning material affects learners' emotional states, and more specifically, that the perceived aesthetic of the learning material affords more positive emotional states in learners. The perceived classical aesthetic even seems to slightly lower negative activation (hypothesis 1.1). Furthermore, the perceived usability marginally reduces negative activation, but affords more positive emotional states (hypothesis 1.2).

6.3.3. Research question 2: regression of the emotional states on learning outcomes

For the second stage, we investigated whether participants' emotional states during learning affected their learning outcomes. We conducted single linear regression analyses with the learners' emotional states during learning (mean of PANAVA 2 and 3) as independent variable and the learning outcomes measures (retention, comprehension, transfer) as dependent variables.

The results suggested that the learners' performance on *retention* tasks was significantly predicted by positive activation (β = .196, p < .001) and negative activation (β = −.112, p = .041) as reported by the participants, but not by the valence of the emotional state (F(1, 332) = 1.36, p = .244). Positive activation explained 3.8% of the variance in retention performance (R^2 = .038, F(1, 332) = 13.22, p < .001), while negative activation explained 1.3% of the variance in retention (R^2 = .013, F(1, 332) = 4.21, p = .041). Participants who reported higher levels of positive activation and lower levels of negative activation performed marginally better on retention tasks.

The analyses indicated that the participants' performance on *comprehension* tasks was significantly predicted by positive activation (β = .113, p = .040) but not affected by negative activation (F(1, 330) = 2.74, p = .099) or valence of the emotional state (F(1, 330) = .07, p = .789). Positive activation explained 1.3% of the variance in comprehension (R^2 = .013, F(1, 330) = 4.26, p = .040). Learners who reported higher levels of positive activation showed slightly better results on comprehension tasks.

The results showed that the learners' performance on *transfer* tasks was significantly predicted by negative activation (β = −.113, p = .040) but not by positive activation (F(1, 330) = 2.60, p = .108) or valence of the emotional state (F(1, 330) = .68, p = .409). Negative activation explained 1.3% of the variance in transfer (R^2 = .013, F(1, 330) = 4.25, p = .040). Hence, participants who showed lower levels of negative activation also showed a marginally better performance on transfer tasks. The results of the regression analyses are visualized in Fig. 3.

6.3.4. Interim discussion

Summing up, positive and negative activation as dimensions of the learners' emotional state predicted performance on retention tasks, albeit in a very subtle way. Furthermore, positive activation very slightly affected comprehension and negative activation very slightly affected transfer performance. However, the valence of the emotional state seems to play a minor role, as valence did not predict any of the learning outcome measures. Taken together, the results support the idea that learners' emotional states during learning affect learning outcomes – although the variance explained by the emotional states is very small. However, the more specific assumptions about the relationships between positive and negative emotional states and the different learning outcome measures are only partly supported. In line with our expectations, the results showed that a more positive emotional state fosters more complex learning goals such as comprehension and transfer (hypothesis 2.1). We further expected that negative emotional states foster retention. Contrary to our expectation, we found that retention performance was facilitated by more positive and less

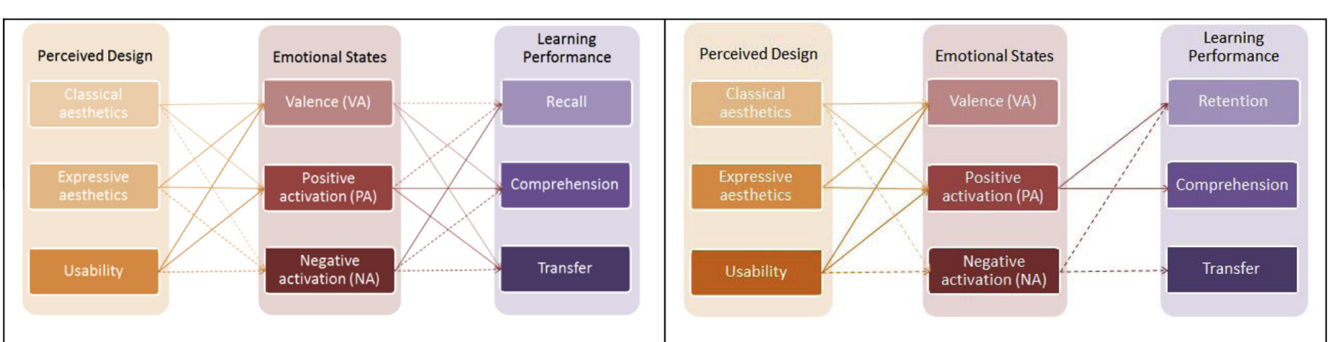

Fig. 3. Expected relationships between perceived design, emotional states and learning outcomes (left) as well as the relationships supported by the data as found in the regression analyses (right). The dotted arrows represent a negative effect.

negative activation (hypothesis 2.2). Taken together, our results showed that positive emotional states facilitated learning outcomes, whereas negative emotional states did not have the expected positive effect but instead hindered learning.

6.3.5. Regression of the emotional states on intrinsic motivation

In order to investigate whether the *emotional states* during learning affect learners' intrinsic motivation, we conducted single linear regression analyses with the learners' emotional states during learning (mean of PANAVA 2 and 3) as independent variable and intrinsic motivation as dependent variable. The results suggest that the learners' intrinsic motivation was significantly predicted by the participants' emotional states during learning (valence: $\beta = .348$, $p < .001$, positive activation: $\beta = .541$, $p < .001$, and negative activation: $\beta = -.149$, $p = .006$). The valence of the learners' emotional state explained 12.1% of the variance in intrinsic motivation ($R^2 = .121$, $F(1, 332) = 45.67$, $p < .001$), positive activation explained 29.3% of the variance ($R^2 = .293$, $F(1, 332) = 137.69$, $p < .001$) and negative activation explained 2.2% of the variance ($R^2 = .022$, $F(1, 332) = 7.50$, $p = .006$). Therefore, in line with our expectations, learners who reported more positive emotional states also showed higher intrinsic motivation. While the negative activation only had a very small effect, the valence of the emotional states and especially positive activation substantially affected the participants' intrinsic motivation during learning (hypotheses 2.3 and 2.4). Intrinsic motivation as assessed in our study not only referred to whether the learners perceived learning with the material as fun and whether they found it interesting, but also asked whether the participants would like to continue working with the material. The results of the regression analyses on intrinsic motivation are visualized in Fig. 4.

7. General discussion

7.1. Do intrinsic design features affect learners' emotional states?

The primary aim of this study was to investigate whether intrinsic design features of multimedia learning material affect learners' emotional states. Based on the conducted pre-study, different color designs were found that are perceived as high vs. low in classical or expressive aesthetics. In the main study, the empirical findings showed that the manipulations of the two aesthetics conditions (classical and expressive) were unsuccessful. Contrary to our expectations, the manipulations of the usability condition (manipulating the loading time) were also unsuccessful. Although the loading time in the low usability condition was distinctly higher than that of the high usability group, no significant differences between the experimental conditions were found in perceived usability. Considering this constraint, the reported results of the between-group comparisons can be interpreted as follows: the objectively provided differences in aesthetics and usability did not affect the perceived aesthetics and usability as rated by the participants. They also did not affect the learners' emotional states during learning. However, regression analyses revealed that the perceived aesthetics and usability affected the learners' emotional states in the expected way. High expressive and classical aesthetics along with good usability facilitated more positive emotional states via positive activation and valence, and the latter two lowered negative activation. In line with our expectations, the results indicate that intrinsic design features of multimedia learning material seem to affect learners' emotional states. In reference to current research (e.g., Plass et al., 2014), we also provided a new viewpoint on emotional design in multimedia learning compared to the seductive details approach (Harp & Mayer, 1997; Mayer et al., 2001), which has critically discussed the integration of appealing but interesting design elements in multimedia learning materials. Our study provides evidence that an emotional design of multimedia learning materials does not necessarily need additional elements, but can instead apply intrinsic design elements such as colors or loading times. Nevertheless, the crucial factor is not the objectively given differences in the design but rather the subjectively perceived levels of aesthetics and usability. Unfortunately, even with a thorough pre-study and systematic deduction of promising emotional design features, the study failed to show which design features might be effective in order to be perceived as highly aesthetic and/or usable. Various factors might explain why the chosen color designs and loading times had no direct effects on perceived aesthetics and usability.

The learning situation may provide one possible explanation for the failed manipulation check. Regardless of their performance, the students in our study received a cinema voucher for taking part in the experiment. Since a high learning test score did not matter, the incentive to learn as well as the valence of performance were low (cf. Pekrun et al., 2011). When the learning performance does matter, e.g. in a real achievement situation, the results might be different (cf. Bessiere et al., 2006; Ceaparu et al., 2002; Lazar et al., 2006). If students want to learn efficiently (e.g., under time pressure), then it is more likely that they will perceive the long waiting times as disturbing. Thus, effects on the negative emotional state of learners – induced by the manipulated loading times – can be expected in real achievement situations. Existing studies suggest that whether loading times are expected for the user also plays a role. When longer loading times can be anticipated, the user tolerates them (Bouch et al., 2000; Dellaert & Kahn, 1999). Although we did not provide direct information on the length of the loading time in our study, the students were probably able to infer the waiting period, as each of the 60 pages of the learning material needed the same time to load. Since the study was conducted in

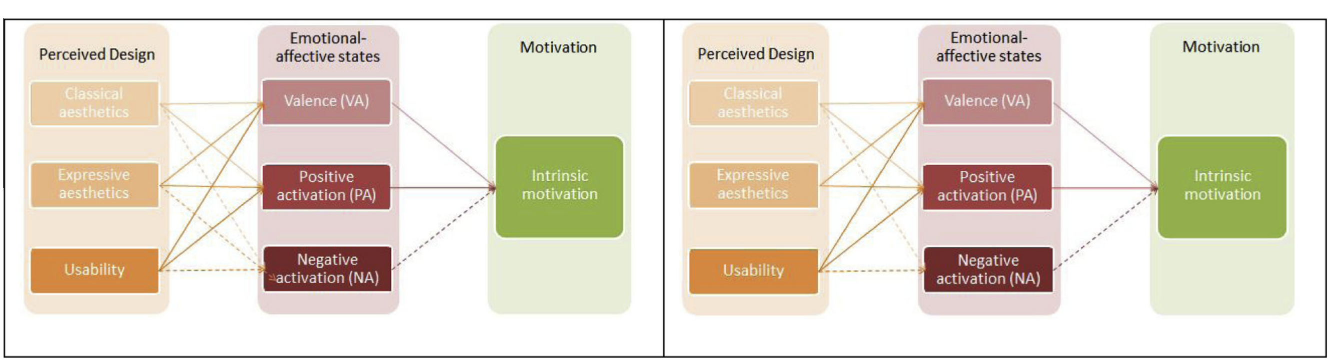

Fig. 4. Expected relationships between perceived design, emotional states and intrinsic motivation (left) as well as the relationships supported by the data as found in the regression analyses (right).

a computer lab at the university, it is also possible that the students were expecting very slow loadings times (e.g., Dellaert & Kahn, 1999). Assuming that frustration is highly dependent on the expectations of the user, future studies are necessary to examine these effects in more detail. Moreover, the subject covered by the learning material had no domain-specific relevance for the target group; it did not correspond to a real learning situation, and this is also a threat to generalization. The participants received a multimedia computer-based program concerning how weather works. Perhaps the chosen topic included concepts which were too simple, and the results might differ for other issues such as more relevant learning material. Therefore, further research is also necessary to investigate the effects in a "real" achievement situation with performance-based incentives and relevant topics.

Another possible explanation for the results might be that people may have lower expectations for the design of multimedia learning materials than for websites. User expectations, however, are a strong determinant of the perceived quality of the material and the experienced level of frustration (e.g., Bessiere et al., 2006; Ceaparu et al., 2002). While we are used to expertly designed, colorful and appealing websites, our experience with learning material might be quite different –black and white copies or poorly readable scans are common on learning platforms. Our chosen target group (college students) may be particularly accustomed to learning with badly designed learning material. College students as target group could be considered as high achieving academic learners; the generalization of the findings is therefore limited to this sample. Moreover, the sample was predominantly female, compared to the more balanced sample of male and female participants in the study by Um et al. (2012). Current research by Plass et al. (2014) also noted that learner variables such as gender, cultural context, and prior knowledge should be given more consideration to determine how specific emotions can be induced through emotional design. Hence, future studies are needed in order to test the hypotheses of our study with a broader target group (e.g., participants of continuing education or high school students, balanced sample of females and males). Further studies are also needed in order to shed light on the assumption of whether individuals may have lower standards for the design of learning material than for websites.

Thirdly, the empirical findings of both the pre-study and the main study should be discussed to explain the failed manipulation check. In the conducted pre-study, the students were able to rate different designs for the learning material on classical and expressive aesthetics. Our main study aimed to compare differences between the experimental conditions and there were accordingly no comparisons to other designs for the learning material (between-subject design). Moreover, the ratings of the perceived aesthetics and usability were based on first impressions in the pre-study, whereas in the main study the different designs were applied in a learning situation. Taken together, the results of the two studies provide several suggestions that can be used to produce a systematic derivation of design criteria for multimedia learning material. To draw general conclusions on how multimedia learning materials should be designed in order to be perceived as aesthetically pleasing, further research is necessary. Further experiments may use a within-subject design, presenting good/bad design in a learning situation to the same students. Combining the experimental methods with interviews might also be useful to gain more insights into how the learners perceive the learning material and to obtain more detailed information about the participants' perceptions. Doing so, could help discover whether the learners perceive the learning material as highly aesthetic and/or usable and also why they feel this way.

In summary, we took concepts from web design into account to systematically deduce possible emotional design features and to take a multidisciplinary view on the research question. In contrast to previous studies on emotional design in multimedia learning (Plass et al., 2014; Um et al., 2012), we systematically deduced intrinsic emotional design features. These were successful in the pre-study (competitive comparison based on first impressions), but unfortunately not in a learning situation as tested in the main study. In extension to existing studies on emotional design (Plass et al., 2014; Um et al., 2012), we used more differentiated measures of the learners' emotional states by applying the PANAVA-KS (Schallberger, 2005), which unlike the PANAS, which solely measures positive and negative activation, also measures valence as an extra dimension (Tellegen et al., 1999). This also adds to conceptual clarity, as the scales of the PANAS were misleadingly referred to as positive and negative affect in previous studies, rather than positive and negative activation (cf. Tellegen et al., 1999). The different results for our study confirmed that we are only able to answer the first research question via a multifaceted approach. Thus, an important theoretical and methodical implication is that a more differentiated view on intrinsic design features of multimedia learning material and on the assessment of learners' emotional states – as attempted in our study – should be retained in future research.

7.2. How do emotional states affect the learning process?

The second aim of this study was to examine how emotional states experienced during learning affect the learning process. This question refers to a current discussion in multimedia learning research: on the one hand, it is assumed that emotional states have an inhibiting effect, because they may place additional demand on the limited resources of working memory and cause task-irrelevant thinking (e.g., Seibert & Ellis, 1991). On the other hand, they may positively affect working memory processes and information retrieval (e.g., Erez & Isen, 2002). Even though the facilitation hypothesis is prevalent in positive emotions research, the influence of learners' emotional states – especially of negative emotional states – on complex learning is not yet clear.

Our study provides further evidence that emotional states can facilitate complex learning processes; however, it is too early to draw general conclusions. As a limitation of the study, it should be noted that the second research question could not be answered by the means of between-group comparisons, as the objectively given differences in design and usability failed to affect the perceived aesthetics and usability (failed manipulation check) and also failed to internally induce the intended emotional states. Nevertheless, the applied regression analyses revealed that the perceived aesthetics and usability did affect the emotional states of the learners in the expected way. The resulting emotional states did then affect the learning outcomes – albeit in a very subtle way. Here, our research profited from taking up a more differentiated perspective on emotional diversity, as we considered positive and negative activation and additionally their valence as dimensions of learners' emotional state. This allowed differentiated statements about the effects on learning: positive and negative activation but not valence slightly affected the learning outcomes. Positive activation subtly affected retention and comprehension, and negative activation slightly affected retention and transfer performance. The valence of the emotional state seems to play a minor role, as valence did not predict one of the learning outcome measures. Taken together, our more specific hypotheses on the relationship between positive and negative emotional states and the different learning outcomes measures are only partly supported: contrary to our expectations, we found that retention performance was facilitated by more positive and less negative activation. In line with our expectations, the results showed that a more positive emotional state fosters more complex learning goals (comprehension and transfer). To explain these findings, the expected effects might be moderated by other variables. In our study, we focused on emotional factors and learning outcomes excluding cognitive factors such as cognitive load. Previous research, however, suggested that positive emotions can increase cognitive load in working memory (e.g., Um et al., 2012). Further investigations should therefore consider these possible moderator variables to analyze their influence on the expected relationship. Bearing in mind that our results showed only weak direct effects of the emotional states on learning, this study still went on to examine the impact of emotional aspects on complex learning processes in more detail. In contrast to the existing studies on emotional design (e.g., Plass et al., 2014; Um et al., 2012), we were able to derive more differentiated hypotheses for the relation between learners' emotional states and learning outcomes, as we did not only consider positive emotional states but also investigated the effect of negative emotional states. A crucial implication of this study is that learners' emotional states do not generally promote the learning process; instead, the impact of positive and negative emotional states depends on the requirements and learning goals (e.g., analytical vs. holistic). Since the different empirical findings showed that emotional state has no straightforward effect on any reasonable learning outcomes measure (retention, comprehension, transfer), future studies should also apply a variety of performance measures in order to reveal the multi-level diversity of effects.

Our study further revealed that a learner's intrinsic motivation is an important variable. In line with previous research on emotional design in multimedia learning (Plass et al., 2014; Um et al., 2012) and with theoretical based models (cognitive-motivational mediation model: Pekrun, 1992), we showed that emotional states during learning quite substantially affected learners' intrinsic motivation – including the motivation to continue working with the material. Although we could only find very subtle direct effects of the induced emotional states on learning outcomes, we found indirect effects on the motivation to spend more time on the material. Hence, when – in an actual learning setting – more time is deliberately spent on the same material, the learning outcomes might be facilitated indirectly. This very important implication of the study provides a number of starting points for future research. For instance, studies with unlimited learning time and "real" learning situations are necessary to investigate motivational effects in more detail. In contrast to most research on multimedia learning that used very short learning material (e.g., Mayer, 2009), the learners in our study worked with the learning material for an average of 30 min. When examining how the length of instruction can moderate learning outcomes, the effects should be tested in various courses for different periods of time. This would allow further insights into whether emotional effects on learning only occur in limited learning settings (e.g., experimental situations), or whether they also occur in "real" achievement situations. Possible long-term effects on learners' emotional states, motivation, and learning could be also analyzed by studies including online courses over a full semester or delayed instead of immediately testing learning performance (cf. Schweppe & Rummer, 2012, subm.). In reference to previous research on emotional design, we also focused on intrinsic motivation; however, learners' motivation – just like learners' emotional states – has different dimensions. Future research should therefore consider additional aspects of learning motivation (e.g. fear of failure) to establish causality between an appealing design of the learning material, learners' motivation, and learning outcomes. Bearing in mind that we focused only on self-reported measurements of emotion and motivation, further studies should also consider additional objective measurements.

In summary, whereas previous studies provided initial insights into the role of positive emotions on learning, we seek to broaden the picture by not only looking at positive emotional states, but also investigating the effect of negative emotional states on complex learning processes. According to the empirical findings, we propose future studies should instead focus on learning effectiveness, rather than solely on efficiency, e.g. by excluding time constraints, including emotional and motivational aspects as well as long-term effects. An important theoretical and practical implication of the present study is that learners' emotional states have no straightforward effects on learning outcomes; instead the effects are traced back to various factors. Our study provides important factors for investigating the interplay between learners' emotional states (positive and negative), learning motivation, and learning outcomes in greater depth. However, further work is inevitable to analyze multidimensional circumstances under which emotional design effects occur in multimedia learning.

7.3. Conclusion

Our research has begun to investigate the questions of whether the design of multimedia learning material affects learners' emotional states, and how these emotions affect complex learning processes. The investigations can be seen as a response to current discussions in multimedia learning research (e.g., Um et al., 2012). The experimental design of our study was generally in line with the design of the previous studies into emotional design in multimedia learning (Plass et al., 2014; Um et al., 2012). Whereas the previous studies successfully provided initial evidence that an emotional design matters in terms of emotion induction and learning effectiveness, the current study aimed to produce a more differentiated view on these effects. We first systematically derived possible emotional design features, refraining from external emotion induction (due to possible confounds) and looking at positive and negative emotional states and their effects on different learning goals (retention, comprehension, transfer). Thus, the results revealed a much more complex picture than the previous studies. In conclusion, a differentiation of the possible effects is necessary, but also requires more differentiated research designs. Future studies into emotional design in multimedia learning should therefore separate the two research questions. On the one hand, studies are needed that identify intrinsic emotional design features that are successful in inducing positive and negative emotional states in learners. On the other hand, studies focusing on the effects of positive and negative emotions on complex learning may for now externally induce the desired emotional states and investigate their effects. A combined approach of internally inducing emotional states and studying their effects may be the next step.

Acknowledgements

We would like to thank the German Science Foundation (DFG) for the support of this work through Project DO 1237/3-1. We would further like to thank Ralf Rummer, Judith Schweppe, and Helmut Niegemann for their valuable comments during the research process and on earlier versions of this manuscript.

References

Ainley, M. (2006). Connecting with learning: Motivation, affect, and cognition. *Educational Psychology Review, 18*, 391–405.

Artino, A. R. (2012). Emotions in online learning environments: Introduction to the special issue. *The Internet and Higher Education, 15*(3), 137–140.

Ashby, F. G., Isen, A. M., & Turken, A. (1999). A neuropsychological theory of positive affect and its influence on cognition. *Psychological Review, 106*, 529–550.

Bagozzi, R. P., Gopinath, M., & Nyer, P. U. (1999). The role of emotions in marketing. *Journal of the Academy of Marketing Science, 27*(2), 184–206.

Bessiere, K., Newhagen, J. E., Robinson, J. P., & Shneiderman, B. (2006). A model for computer frustration: The role of instrumental and dispositional factors on incident, session, and post-session frustration and mood. *Computers in Human Behavior, 22*, 941–961.

Bless, H. T., & Igou, E. R. (2006). Stimmung und Informationsverarbeitung. In H. W. Bierhoff & D. Frey (Eds.), *Handbuch der psychologie: Sozialpsychologie und kommunikationspsychologie* (pp. 423–429). Göttingen: Hogrefe.
Bless, H. (2001). The consequences of mood on the processing of social information. In A. Tesser & N. Schwarz (Eds.), *Blackwell handbook in social psychology* (pp. 391–412). Oxford, UK: Blackwell Publishers.
Blythe, M., & Wright, P. (2003). Introduction – from usability to enjoyment. In M. Blythe, K. Overbeeke, A. Monk, & P. Wright (Eds.), *Funology: From usability to enjoyment* (pp. XIII–XIX). Dordrecht: Kluwer.
Bouch, A., Bhatti, N., & Kuchinsky, A. J. (2000). Quality is in the eye of the beholder: Meeting users' requirements for internet quality of service. In: *Proceedings of the CHI'2000* (pp. 297–304), The Hague, The Netherlands.
Bradley, M. M., & Lang, P. J. (1994). Measuring emotion: The self-assessment manikin and the semantic differential. *Journal of Behavior Therapy und Experimental Psychiatry, 25*, 49–59.
Cacioppo, J. T., Gardner, W. L., & Berntson, G. G. (1999). The affect system has parallel and integrative processing components: Form follows function. *Journal of Personality and Social Psychology, 76*, 839–855.
Ceaparu, I., Lazar, J., Bessiere, K., Robinson, J., & Shneiderman, B. (2002). Determining causes and severity of end-user frustration. *International Journal of Human–Computer-Interaction, 17*, 333–356.
Chanlin, L. (1998). Animation to teach students of different knowledge levels. *Journal of Instructional Psychology, 25*(3), 166–175.
Chen, C. M., & Wang, H. P. (2011). Using emotion recognition technology to assess the effects of different multimedia materials on learning emotion and performance. *Library and Information Science Research, 33*, 244–255.
Coursaris, C. K., Swierenga, S. J., & Watrall, E. (2008). An empirical investigation of color temperature and gender effects on web aesthetics. *Journal of Usability Studies, 3*(3), 103–117.
Dellaert, B. G. C., & Kahn, B. E. (1999). How tolerable is delay? Consumer's evaluations of internet web sites after waiting. *Journal of Interactive Marketing, 13*, 41–54.
Elliot, A. J., Maier, M. A., Binser, M. J., Friedman, R., & Pekrun, R. (2009). The effect of red on avoidance behavior. *Personality and Social Psychology Bulletin, 35*, 365–375.
Erez, A., & Isen, A. M. (2002). The influence of positive affect on components of expectancy motivation. *Journal of Applied Psychology, 87*(6), 1055–1067.
Fiedler, K. (1988). Emotional mood, cognitive style, and behavior regulation. In K. Fiedler & J. Forgas (Eds.), *Affect, cognition and social behavior* (pp. 100–119). Toronto: Hogrefe & Huber.
Fredrickson, B. L. (2003). The value of positive emotions. *American Scientist, 91*, 330–335.
Frijda, N. H. (1999). Emotions and hedonic experience. In D. Kahneman, E. Diener, & N. Schwarz (Eds.), *Well-being: Foundations of hedonic psychology* (pp. 190–210). New York: Russell Sage Foundation.
Gable, S. L., Reis, H. T., & Elliot, A. J. (2003). Evidence for bivariate systems: An empirical test of appetition and aversion across domains. *Journal of Research in Personality, 37*, 349–372.
Galletta, D., Henry, R., McCoy, S., & Polak, P. (2002). Web site delays: How slow can you go? In: *Paper presented at the first annual pre-ICIS workshop on HCI research in MIS*, Barcelona, Spain.
Garcia, L., Rogat, T. K., & Koskey, K. L. K. (2011). Affect and engagement during small group instruction. *Contemporary Educational Psychology, 36*, 13–24.
Garner, R., Gillingham, M. G., & White, C. S. (1989). Effect of "seductive details" on macroprocessing and microprocessing in adults and children. *Cognition and Instruction, 6*, 41–57.
Goetz, T., Nett, U. E., Martiny, S. E., Hall, N. C., Pekrun, R., Dettmers, S., et al. (2012). Students' emotions during homework: Structures, self-concept antecedents, and achievement outcomes. *Learning and Individual Difference, 22*(2), 225–234.
Goetz, T., Pekrun, R., Hall, N., & Haag, L. (2006). Academic emotions from a social-cognitive perspective: Antecedents and domain specificity of students' affect in the context of Latin instruction. *British Journal of Educational Psychology, 76*(2), 289–308.
Harp, S. F., & Mayer, R. E. (1997). The role of interest in learning from scientific text and illustrations: On the distinction between emotional interest and cognitive interest. *Journal of Educational Psychology, 89*(1), 92–102.
Hartmann, J., Sutcliffe, A., & De Angeli, A. (2007). Investigating attractiveness in web user interfaces. In *Proceedings of the SIGCHI conference on human factors in computing systems (CHI'07)* (pp. 387–396). New York, NY, USA: ACM.
Hassenzahl, M. (2004). Beautiful objects as an extension of the Self. A reply. *Human Computer Interaction, 19*, 377–386.
Hassenzahl, M., Burmester, M., & Beu, A. (2001). Engineering joy. *IEEE Software, 1 and 2*, 70–76.
Hassenzahl, M., & Tractinsky, N. (2006). Empirical studies of the user experience. *Behavior and Information Technology, 25*(2), 91–97 [special issue].
Isen, A. M. (1984). Toward understanding the role of affect in cognition. In R. Wyer & T. Srull (Eds.), *Handbook of social cognition* (pp. 174–236). Hillsdale, NJ: Erlbaum.
Isen, A. M., Johnson, M. M. S., Mertz, E., & Robinson, G. F. (1985). The influence of positive affect on the unusualness of word associations. *Journal of Personality and Social Psychology, 48*, 1–14.
Isen, A. M. (2000). Positive affect and decision making. In M. Lewis & J. M. Haviland (Eds.), *Handbook of emotions* (2. Aufl., pp. 417–435). New York: Guilford Press.
Isen, A. M., & Reeve, J. (2005). The influence of positive affect on intrinsic and extrinsic motivation: Facilitating enjoyment of play, responsible work behaviour, and self-control. *Motivation and Emotion, 29*(4), 295–323.
Itten, J. (1989). *Kunst der Farbe*. Ravensburg: Maier.
Kaya, N., & Epps, H. (2004). Relationship between color and emotion: A study of college students. *College Student Journal, 38*(3), 396–405.
Lavie, T., & Tractinsky, N. (2004). Assessing dimensions of perceived visual aesthetics of websites. *International Journal of Human-Computer Studies, 60*(3), 269–298.
Lazar, J., Jones, A., Hackley, M., & Shneiderman, B. (2006). Severity and impact of computer user frustration: A comparison of student and workplace users. *Interacting with Computers, 18*(2), 187–207.
Lehman, S., Shraw, G., McCrudden, M., & Hartley, K. (2007). Processing and recall of seductive details in scientific text. *Contemporary Educational Psychology, 32*, 569–587.
Lenzner, A. (2009). *Visuelle Wissenskommunikation: Effekte von Bildern beim Lernen. Kognitive, affektive und motivationale effekte*. Hamburg: Kovac.
Leutner, D. (2014). Motivation and emotion as mediators in multimedia learning. *Learning and Instruction*. Advance online publication. http://dx.doi.org/10.1016/j.learninstruc.2013.05.004.
Liedl, R. (1994). *Die Pracht der Farben*. Mannheim: BI Wissenschaftsverlag.
Linnenbrink, E. A. (2006). Emotion research in education: Theoretical and methodological perspectives on the integration of affect, motivation, and cognition. *Educational Psychology Review* (advance online publication DOI:http://dx.doi.org/10.1007/s10648-006-9028-x.).
Mayer, R. E. (2009). *Multimedia learning* (2. Aufl.). Cambridge: Cambridge University Press.
Mayer, R. E., Heiser, H., & Lonn, S. (2001). Cognitive constraints on multimedia learning: When presenting more material results in less understanding. *Journal of Educational Psychology, 93*, 187–198.
Mayer, R. E., & Moreno, R. (2003). Nine ways to reduce cognitive load in multimedia learning. *Educational Psychologist, 38*, 43–52.
Mentis, H. M., & Gay, G. K. (2003). User recalled occurrences of usability errors: Implications on the user experience. In: *Proceedings CHI EA'03. Extended abstracts on human factors in computer systems* (pp. 736–737).
Meyer, D. K., & Turner, J. C. (2007). Scaffolding emotions in classrooms. In P. A. Schultz, R. Pekrun, & P. Winne (Eds.), *Handbook of educational psychology* (pp. 327–348). Mahwah, NJ: Erlbaum.
Moshagen, M., Musch, J., & Göritz, A. S. (2009). A blessing, not a curse: Experimental evidence for beneficial effects of visual aesthetics on performance. *Ergonomics, 52*(10), 1311–1320.
Moshagen, M., & Thielsch, M. T. (2010). Facets of visual aesthetics. *International Journal of Human–Computer Studies, 68*, 689–709.
Müller, J. (2009). *Joy of use: Die messung von emotionen und ästhetik auf websites*. Erziehungswissenschaftliche Fakultät der Universität Erfurt, Unveröffentlichte Diplomarbeit.
Müller, J., Heidig, S., & Niegemann, H. M. (2012). Evoking emotional dimensions in HCI – development of the questionnaire user experience (QUX). In: *Paper presented at the annual meeting of the American educational research association (AERA)*, April 13th–17th, Vancouver, Canada.
Müller, J., Heidig, S., & Reichelt, M. (subm.). Color matters: Visual aesthetics and their emotional-affective impact in multimedia learning environments. *Journal of Media Psychology*.
Nah, F. (2004). A study on tolerable waiting time: How long are Web users willing to wait? *Behaviour and Information Technology, 23*(3), 153–163.
Nielsen, J. (1997). *The need for speed. Jakob Nielsen's Alertbox for March, 1997*. Available at <http://www.useit.com/alertbox/9703a.html>.
Norman, D. A. (2002). Emotion and design: Attractive things work better. *Interactions Magazine, 4*(4), 36–42.
Norman, D. A. (2004). Introduction to this special section on beauty, goodness, and usability. *Human Computer Interaction, 19*, 311–318.
Oaksford, M., Morris, F., Grainger, B., & Williams, J. M. G. (1996). Mood, reasoning, and central executive processes. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 22*(2), 476–492.
Paas, F., Renkl, A., & Sweller, J. (2003). Cognitive load theory and instructional design: Recent developments. *Educational Psychologist, 38*, 1–4.
Park, B., Plass, J.L., & Brünken, R. (2014). *Cognitive and affective processes in multimedia learning. Learning and Instruction* (advance online publication <http://dx.doi.org/10.1016/j.learninstruc.2013.05.005>).
Park, S., & Lim, J. (2007). Promoting positive emotion in multimedia learning using visual illustrations. *Journal of Educational Multimedia and Hypermedia, 16*(2), 141–162.
Pekrun, R. (1992). The impact of emotions on learning and achievement: Towards a theory of cognitive/motivational mediators. *Applied Psychology: An International Review, 41*, 359–376.
Pekrun, R. (2006). The control-value theory of achievement emotions: assumptions, corollaries, and implications for educational research and practice. *Educational Psychology Review* (advance online publication: doi:http://dx.doi.org/10.1007/s10648-006-9029-9).
Pekrun, R., Elliot, A. J., & Maier, M. A. (2009). Achievement goals and achievement emotions: Testing a model of their joint relations with academic performance. *Journal of Educational Psychology, 101*, 115–135.
Pekrun, R., Goetz, T., Titz, W., & Perry, R. P. (2002). Academic emotions in students' self-regulated learning and achievement: A program of qualitative and quantitative research. *Educational Psychologist, 37*(2), 91–105.
Pekrun, R., Götz, T., Frenzel, A. C., Barchfeld, P., & Perry, R. P. (2011). Measuring emotions in students' learning and performance: The achievement emotions questionnaire (AEQ). *Contemporary Educational Psychology, 36*, 36–48.
Pintrich, P. R. (2000). Multiple goals, multiple pathways: The role of goal orientation in learning and achievement. *Journal of Educational Psychology, 92*, 544–555.

Plass, J. L., Heidig, S., Hayward, E. O., Homer, B. D., & Um, E. (2014). Emotional design in multimedia learning: Effects of shape and color on affect and learning. *Learning and Instruction, 29*, 128–140.
Plass, J. L., Moreno, R., & Brünken, R. (2010). *Cognitive load theory*. Cambridge: Cambridge University Press.
Ramsay, J., Barbesi, A., & Preece, J. (1998). A psychological investigation of long retrieval times on the world wide web. *Interacting with Computers, 10*, 77–86.
Reeps, I. E. (2006). *Joy-of-Use – Ästhetik, Emotion und User Experience für interaktive Produkte*. VDM.
Rey, G. D. (2012). A review of research and a meta-analysis of the seductive detail effect. *Educational Research Review, 7*, 216–237.
Rosenberg, E. L. (1998). Levels of analysis and the organization of affect. *Review of General Psychology, 2*, 247–270.
Russell, J. A. (1980). A circumplex model of affect. *Journal of Personality and Social Psychology, 39*, 1161–1178.
Russell, J. A. (2003). Core affect and the psychological construction of emotion. *Psychological Review, 110*, 145–172.
Schallberger, U. (2005). Kurzskalen zur Erfassung der Positiven Aktivierung, Negativen Aktivierung und Valenz in Experience Sampling Studien (PANAVA-KS). Theoretische und methodische Grundlagen, Konstruktvalidität und psychometrische Eigenschaften bei der Beschreibung intra- und interindividueller Unterschiede. Forschungsberichte aus dem Projekt: ≪Qualität des Erlebens in Arbeit und Freizeit≫, Nr. 6. Zürich: Fachrichtung Angewandte Psychologie des Psychologischen Instituts der Universität.
Schenkmann, B. N., & Jönsson, F. U. (2000). Aesthetics and preferences of web pages. *Behavior and Information Technology, 19*(5), 367–377.
Schwarz, N. (1990). Feelings as information: Informational and motivational functions of affective states. In E. T. Higgins & R. Sorrentino (Eds.). *Handbook of motivation and cognition: Foundations of social behavior* (vol. 2, pp. 527–561). New York: Guilford Press.
Schwarz, N., & Clore, G. L. (1996). Feelings and phenomenal experiences. In E. T. Higgins & A. Kruglanski (Eds.), *Social psychology: Handbook of basic principles* (pp. 433–465). New York: Guilford.
Schweppe, J., & Rummer, R. (subm.). *Written text presentation as a desirable difficulty in long-term multimedia learning* (submitted for publication).
Schweppe, J., & Rummer, R. (2012). Long-term multimedia learning. In E. de Vries & K. Scheiter (Eds.), *Proceedings EARLI special interest group text and graphics: Staging knowledge and experience: How to take advantage of representational technologies in education and training?* (pp 196–198). Grenoble, France: Université Pierre-Mendès-France.
Seibert, P. S., & Ellis, H. C. (1991). Irrelevant thoughts, emotional mood states, and cognitive task performance. *Memory and Cognition, 19*, 507–513.
Selvidge, P. R., Chaparro, B., & Bender, G. T., (2000). The world wide wait: Effects of delays on user performance. In: *Proceedings of the IEA 2000/HFES 2000 Congress* (pp. 1-416–1-419).
Tellegen, A., Watson, D., & Clark, L. A. (1999). On the dimensional and hierarchical structure of affect. *Psychological Science, 10*, 297–303.
Thielsch, M. T. (2008). *Ästhetik von Websites: Wahrnehmung von Ästhetik und deren Beziehung zu Inhalt, Usability und Persönlichkeitsmerkmalen*. Münster: MV Wissenschaft.
Tractinsky, N., Shoval-Katz, A., & Ikar, D. (2000). What is beautiful is usable. *Interacting with Computers, 13*(2), 127–145.
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.
Um, E., Plass, J. L., Hayward, E. O., & Homer, B. D. (2012). Emotional design in multimedia learning. *Journal of Educational Psychology, 104*(2), 485–498.
Valdez, P., & Mehrabian, A. (1994). Effects of color on emotion. *Journal of Experimental Psychology, 123*(4), 394–409.
Van der Heijden, H. (2003). Factors influencing the usage of websites: the case of a generic portal in the Netherlands. *Information and Management, 40*(6), 541–549.
van Schaik, P., & Ling, J. (2009). The role of context in perceptions of the aesthetics of web pages over time. *International Journal of Human-Computer Studies, 67*, 79–89.
Watson, D., Clark, L. A., & Tellegen, A. (1988). Development and validation of brief measures of positive and negative affect: The PANAS scales. *Journal of Personality and Social Psychology, 54*, 1063–1070.
Watson, D., & Tellegen, A. (1985). Toward a consensual structure of mood. *Psychological Bulletin, 98*, 219–235.
Weller, L., & Livingston, R. (1988). Effect of colour of questionnaire on emotional responses. *Journal of General Psychology, 115*, 433–440.
Zhang, Q., & Lee, M. (2009). Analysis of positive and negative emotions in natural scene using brain activity and GIST. *Neurocomputing, 72*, 1302–1306.