Computers and Education Open 3 (2022) 100075

Contents lists available at ScienceDirect

Computers and Education Open

journal homepage: www.sciencedirect.com/journal/computers-and-education-open

Aesthetic design of app interfaces and their impact on secondary students' interest and learning

Alessia Ruf [*,a,d], Carmen Zahn [a], Doris Agotai [b], Glena Iten [c], Klaus Opwis [d]

[a] University of Applied Sciences and Arts Northwestern Switzerland, School of Applied Psychology, Switzerland
[b] University of Applied Sciences and Arts Northwestern Switzerland, School of Engineering, Switzerland
[c] The Schwyz University of Teacher Education, Switzerland
[d] University of Basel, Faculty of Psychology, Switzerland

ARTICLE INFO

Keywords:
Media in education
Mobile learning
Secondary education
21st century abilities

ABSTRACT

Interest in science topics is an important prerequisite for science learning and achievement. Here, as part of a field experiment, we studied whether teenagers' interest and learning of physics topics would be influenced by the aesthetics of a multimedia learning app. More specifically, we investigated with the example of learning about energy (types of power plants) how different interface designs of a multimedia learning app would influence aesthetic experience, interest, and learning outcome. In our study Swiss high school students (N = 108) were assigned to one of two conditions (i.e., *game-style* vs. *industrial-style*) differing in various aesthetic features. Results indicate that high-quality interfaces support learning and expressive aesthetic design features additionally foster interest in order to engage with the topic. Moreover, our findings on aesthetic experience suggest that *deep perceptual processes*, such as emotion and cognitive stimulation induced by interfaces, further impact interest and learning. Thus, our study gives implications for the design of interest-generating and learning-supporting science apps for teenagers and emphasizes the significance to consider aesthetic experience in future research.

1. Introduction

Interest in science and understanding science topics is key in everyday life to raise awareness and willingness to engage with global issues and current public debates, such as climate change [49,51], and in order to stand up and take responsibility [4]. It is also central for good decision-making based on sensible arguments. For instance, science literacy was proven to significantly decrease the likelihood of people believing in health rumors [21]. However, interest in science topics is steadily declining [50] – science-related topics are often perceived as difficult, boring, and challenging to learn, provoking negative emotions and a lack of interest [5,71]. Thus, finding timely ways to foster positive emotions and interest in science topics is important. If students could learn science topics in ways that are considered fun, this could help increase their overall science literacy and, moreover, in the long run help foster more future science careers [1,58].

One possibility to help cultivate scientific interest and learning is to create interesting computer-supported learning environments, such as multimedia learning apps [10,27]. As interest is a crucial component of intrinsic motivation [25,79], such environments could be seen as a first step in increasing motivation to learn more [e.g., 59], which in turn should increase learning performance [62] and scientific literacy [13].

Yet, what is considered "interesting" for teenage students? Research indicates that interface aesthetics plays an important role in fostering interest, positive emotions, engagement, and enthusiasm in computer-supported classrooms [26,74]. Three recently published meta-analyses provide empirical evidence that appealing aesthetic interfaces of learning environments can have impacts on emotional and interest-related factors and facilitating effects on learning [7,68,78]. These effects particularly occurred when learners interacted with mobile devices and software applications [68] and were stronger in K-12 (primary and secondary students) than in post-secondary students [7,78].

Thus, aesthetic interfaces can enhance interest in science-topics. However, few studies on the effects of aesthetic design were conducted with secondary students [7,78], even though this age group was found to be a critical life period in which orientations in science can

* Corresponding author.
E-mail addresses: alessia.ruf@fhnw.ch (A. Ruf), carmen.zahn@fhnw.ch (C. Zahn), doris.agotai@fhnw.ch (D. Agotai), glena.iten@phsz.ch (G. Iten), klaus.opwis@unibas.ch (K. Opwis).

https://doi.org/10.1016/j.caeo.2022.100075
Received 26 June 2021; Received in revised form 15 January 2022; Accepted 15 January 2022
Available online 17 January 2022

successfully be encouraged [1,58]. The present research uses a field study and measures of learners' aesthetic experience to investigate the influence of two differently designed appealing and learning-supportive app interfaces on secondary school students. In the next section we will first clarify the terms aesthetic design and aesthetic experience in relation to learning. Then, we will provide a summary of related research on emotional design and its effects on scientific interest and learning, before finally emphasizing important research gaps and ending the section with the goals, research questions, and hypotheses of the present study.

1.1. Aesthetic design and aesthetic experience in HCI research

Aesthetic design is a construct that can be described as any design (e.g., websites, interactive systems) that immediately increases the appeal and attractiveness of an object for its observer [e.g., 45,67]. Although the term still lacks a consistent definition [68], the literature provides evidence that single aesthetic features, such as color [e.g., 52,65], familiarity [e.g., 32,57], expressivity [e.g., 30], and visual complexity [e.g., 65,69], have an impact on the perceived attractiveness of an object [22,34] and increase visual appeal, which is "an immediate pleasurable subjective experience that is directed toward an object and not mediated by intervening reasoning" [p. 690, 45]. This is in line with the Model of Aesthetic Art Experience by Leder et al. [31]. In their model, viewers pass through several stages when they contemplate an aesthetic object, such as artwork: perceptual analyses (e.g., complexity of an object), implicit memory integration (e.g., prototypicality of an object), explicit classification (e.g., style of the object), cognitive mastering (e.g., self-related interpretation), and evaluation (e.g., cognitive and affective state). The latter two are closely related in the form of a feedback-loop: successful cognitive mastering leads to a positive evaluation and successful understanding of the aesthetic object that, in turn, positively influences affective states, resulting in pleasure or satisfaction. These experiences of viewers while contemplating an aesthetic object are also closely related to the characteristics that define aesthetic experience as stated by Marković [35]. Marković [35] identified the following aspects as crucial and distinctive: (1) a *motivational, orientational*, or *attentive* aspect of aesthetic experience, defined as a state of intense attention, engagement, or high vigilance of the viewer, (2) a *cognitive* aspect, defined as the viewer's appraisal of an aesthetic object as part of a symbolic (or 'virtual') reality and the transcendence of their everyday uses and meanings, and (3) an *affective* aspect, defined as a strong and clear feeling of unity with aesthetic appraisal.

Research on human-computer interaction (HCI) has applied these definitions to digital interfaces (e.g., websites, online environments) as an important component of user experience [e.g., 11,22] and assumes that aesthetic interfaces are able to improve users' experience of aesthetics [45]. However, there have been few validated instruments measuring viewers' *aesthetic experience* that have been developed so far, and those that have been developed have mainly been for website evaluation [6,30,45,68]. One of the most prominent ones is the Visual Aesthetics of Websites Inventory (VisAWI). It is based on the assumption that users experience the aesthetics of websites according to four facets (i.e., aesthetic features) that are known to have an impact on attractiveness and visual appeal [e.g., 65,69]. The four facets are simplicity, variety, colourfulness, and craftsmanship [45].

Additionally, the advantages of aesthetic design to support learning and task performance have been recognized. This is reflected in the considerable amount of HCI research that examines the impact of a wide range of single aesthetic features [for an overview see 67,68]. In their meta-analysis, Thielsch et al. [68] examined studies that investigated the impact of aesthetic and unaesthetic designs on learning and task performance and found a significant – albeit small ($g = 0.12$) – effect for appealing aesthetic designs. Furthermore, this effect was found to be stronger when users interacted with mobile devices and software applications, which is in keeping with research on computer-supported collaborative learning [10]. However, Thielsch et al. [68] mentioned several limitations of these studies. Many studies did not sufficiently report which aesthetic features (e.g., color, texture etc.) were manipulated and in what way, and used unstandardized or not validated scales for measuring viewers' *aesthetic experience* to assign the design variants to the experimental conditions (aesthetic vs. unaesthetic) or for manipulation check. Hence, Thielsch et al. [68] concluded that it often remained untraceable how appealing or unappealing the designs actually were and how they differed from each other. Yet, positive affective states resulting from a successful aesthetic experience seem to still increase learning performance: in consideration of Norman's Affect Mediation Theory [cf. 46,47], Thielsch et al. [68] argued that aesthetic designs can provoke positive emotions, leading to a better working cognitive system, which results in improved performance. This is in line with the related research field of *emotional design*, which is discussed in the following section.

1.2. Emotional design, learning, and interest in science

Research on *emotional design* is based on the Cognitive Affective Theory of Learning with Media [39,41], which assumes an increase of motivational and affective factors through specific aesthetic features that, in turn, facilitate cognitive processing and, thus, interest and learning [55,56]. Historically, research on emotional design has focused primarily on examining its impact on (natural) science subjects, with limited research in other subjects, such as humanities [76,78]. One reason for this is that these subjects themselves do not usually evoke positive emotions in learners, unlike other subjects, such as poetry, music, or the arts [cf. 76]. Therefore, research on emotional design aims to artificially evoke positive emotions in learners through the provision of emotional design features. Building on the pioneering work of Um et al. [72], the features *color* [e.g., 54,72] and *facial anthropomorphisms in non-human graphical elements* [e.g., 54,64] have been widely investigated [7,78]. In fact, two recent meta-analyses revealed significant effects of emotional designs – compared to neutral designs (i.e., colorless; no facial anthropomorphisms) – on intrinsic motivation, liking/enjoyment, positive affect, and on learning performance [7,78]. Moreover, although the majority of the investigated studies were conducted with post-secondary learners, even stronger effects were found with younger K-12 participants [78]. *Emotional designs* seem to provide an effective way to foster scientific literacy, as they not only facilitate learning but also are able to increase interest in science topics [7,78], which, in turn, could encourage students to pursue a science career in the future.

Since fostering interest is of crucial importance for the present study, we briefly elaborate on the definition and conceptualization of the term in the following: a prominent definition categorizes interest as a critical motivational variable, as well as a psychological state, that influences learning and achievement, which occurs during interactions between persons and their objects of interest [23,24,29]. Thus, interest describes two different experiences: *situational* and *individual interest*. On one hand, *situational interest* is defined as a momentary experience triggered by an object and characterized by increased affect, effort, and attention. In research on science learning, recent work has demonstrated that young learners' *situational interest* in STEM topics could be fostered by providing mobile devices for learning [27]. On the other hand, *individual interest* is described as a more prolonged experience, characterized by a persistent willingness to return to a certain object or topic over time [see 20, for more information]. Theoretical and empirical work suggests that both experiences are closely related [e.g., 24,33]. For example, Romine et al. [62] found that individual interest in science facilitates situational interest. This, in turn, positively impacts science learning. While research on *emotional design* focuses merely on situational interest, both experiences were considered separately in the present study (see section 2.3).

1.3. Current research gaps

It becomes apparent from the theoretical background described above that aesthetic and emotional designs have the ability to positively impact scientific interest and learning [7,68,78]. However, there are still knowledge gaps in the scientific literature that we aim to address in the current study (see section 1.4 for study goals). Firstly, as stated by Thielsch et al. [68], in previous work regarding aesthetic design it is often untraceable how aesthetic or unaesthetic the compared interfaces actually were. This raises the question of whether the effects on learning occurred due to differently manipulated aesthetics between the design variants or due to beneficial or detrimental effects resulting from different implementations of multimedia design principles. Sweller's (66) Cognitive Load Theory and Mayer's (37) Cognitive Theory of Multimedia Learning, two well-established approaches from instructional psychology, incorporate several principles that should be applied when developing computer-supported learning environments to ensure an optimal learning support. The importance of considering these principles when examining the effects of aesthetics on learning is illustrated by the following example: using color – a frequently used design manipulation in research on aesthetic and emotional design [7,68,78] – may, on one hand, induce positive affects in learners, but may, on the other hand, also highlight important parts of the learning material, due to the so-called signaling effect [cf. 36]. In other words: a colorful design (in contrast to a black-and-white design) could not only be beneficial for learning because it triggers higher positive affects, but also because color could function as a learning-supportive instructional design feature.

Secondly, while measurements of viewers' *aesthetic experience* generally remained unconsidered in research on emotional design, it has been embraced by research on aesthetic design (see section 1.1). Nevertheless, in-depth investigations and impacts of the construct on interest and learning are still missing [see section 1.1 and 68]. Hence, only superficial parts of viewers' perceptions provoked by the designs' *surface structures* were investigated, such as perceptions of color or complexity [see VisAWI; 45]. However, in line with earlier approaches on aesthetic art experience, we assume that *aesthetic experience* derives not only due to the perception of *surface structures*, but also due to *deeper perceptual processes* resulting from an active engagement with and exploration of an aesthetic object that might be related to interest and learning [12,31,35]. More precisely, we assume that the experience of aesthetics – especially when it comes to digital interactive environments – is not limited only to perceptions of (static) surfaces of aesthetic interfaces, but is also influenced by the active engagement and interaction with them [see also 8]. Measures considering such *deeper perceptual processes* can be found in related research on general user experience [for an overview see 28] and aesthetic art experience, and are described, for example, as "cognitive stimulation" [19], "intrinsic motivation" [e.g., 25,31], "pleasurable interaction" [30], "learning related boredom" [53], and "involvement" [48]. However, an instrument measuring aesthetic experience by considering both *surface structures* and *deeper processes* of learners who actively engage with aesthetic designs is still missing.

Thirdly, young K-12 learners have rarely been investigated as subjects of research on aesthetic and emotional design [7,68,78] – although findings suggest that effects may be even stronger for younger learners than for post-secondary students [78]. Also, early intervention to spark students' interest is critical, not only in terms of fostering scientific literacy [73], but also because developing science aptitudes in middle school was found to be positively associated with pursuing a science career in the future [1,58].

1.4. Goals of the present study

Consequently, with the present work, we aimed to deepen our understanding of how aesthetic design can foster teenagers' interest and learning in science and how the experience of aesthetic *surface structures* and *deeper perceptual processes* further influence these outcomes.

More precisely, we pursued three main objectives. The first objective was to provide new insights into the effects of aesthetic design on interest and learning in science. In contrast to fundamental research on aesthetic (see section 1.1) and emotional design (see section 1.2), we were less interested in investigating differences of aesthetic vs. unaesthetic or emotional vs. neutral designs than we were in investigating the effects of differently designed appealing interfaces that were equally supportive of learning using systematic and conceptual consideration of multimedia design principles [similar to approaches from instructional psychology, see 37,66]. This new focus allowed us to examine the construct aesthetics (cf. section 1.1) more *holistically* by manipulating multiple aesthetic features in each of the designs. Furthermore, we were able to incorporate principles from the Gestalt Theory [e.g., 3,60] and interpret interfaces as a *Gestalt* that constitutes "a functional unit with properties not derivable by summation of its parts" [p. 187, 3]. Hence, we created two high-quality interface designs of a learning application that oriented towards different life worlds related to the target group (i.e., teenage students) and played with expectations: a *game-style* variant that represented the look and feel of a mobile game and an *industrial-style* variant that was designed in a more technical way and more strongly resembled a classic school book (cf. section 2.2 and Fig. 1, for detailed information). We thereby clearly distinguished between design features that are supportive for learning (i.e., instructional design: committed to be similar between the design variants) and address the target group equally, and appealing aesthetic features that might elicit positive affect in learners but do not directly influence learning due to beneficial or detrimental effects (i.e., *Interface Aesthetic*: permitted to deviate between the design variants, see Table 1). Note that we used physics (more precisely *energy*) as science topic for three main reasons. First, learning physics was found to foster students' scientific literacy through transferring knowledge about scientific products, processes, and attitudes [9,15]; it enables students to make connections between the learned material and their daily lives (e.g., where does the electricity come from?). Second, it involves learning complex processes that can be facilitated by aesthetic designs (see sections 1.1 and 1.2) and by multimedia learning environments as they are able to minimize cognitive load [37]. Third, it has been shown that taking a physics class in high school highly correlates with deciding on a science career in the future [18].

The second objective was to further investigate the effects of aesthetics by examining (explorative) the impact of aesthetic experience on interest and learning performance using a more holistic perspective of the process occurring when a viewer is exposed to and interacts with an aesthetic interface design [31,35]. We therefore investigated the impact of different dimensions of aesthetic experience, including both *surface structures* (i.e., "expressive aesthetics") and *deeper perceptual processes* (i.e., "emotion" and "cognitive stimulation") on interest and learning. For this purpose and by considering measures from closely related research on aesthetic art design and general user experience (see section 1.3), we developed and validated an appropriate instrument at our university that was used in the present study (see section 2.4 for more information).

The third objective was to contribute new results concerning the impact of visually aesthetic designs of a tablet-based learning application on aesthetic experience, interest, and learning of *middle school teenage students* and continue the tradition of research on aesthetic and emotional design [7,68,78]. The study was conducted directly in the classroom.

1.5. Research questions and hypotheses

In order to achieve our study goals (see section 1.4) and fill the above mentioned research gaps (see section 1.3), our study was guided by the following research questions and hypotheses:

Research Question 1. Do the two interface designs (*game-style* vs. *industrial-style*) differ in students' aesthetic experience?

Fig. 1. The figure displays the two interface designs; on the left side the *game-style* design and on the right side the *industrial-style* design (i.e., *Interface Aesthetic*).

Based on research described in sections 1.1 and 1.2, we expected that students learning with the *game-style* design would rate (H1a) *surface structures* (visual aesthetics: VisAWI) and (H1b) *aesthetic experience* (i.e., holistic consideration of *surface structures* and *deeper perceptual processes*) higher than students learning with the *industrial-style* design.

Research Question 2. Do the two interface designs (*game-style* vs. *industrial-style*) differ in their impact on students' situational and individual interest?

Based on interest theories and research on emotional design described in section 1.2, we hypothesized that the *game-style* design would lead to a higher increase in students' (H2a) situational and (H2b) individual interest compared to the *industrial-style* design. Furthermore, we expected that both (H2c) situational and (H2d) individual interest would increase after learning for both designs.

Research Question 3. (explorative) Does aesthetic experience (*surface structures* and *deep processing*) have an impact on situational and individual interest?

According to Research Question 2 and earlier approaches on art experience [31,35], we hypothesized that aesthetic experience (i.e., "expressive aesthetics," "emotion," "cognitive stimulation") and usability would have an impact on (H3a) situational and (H3b) individual interest.

Research Question 4. (explorative) Do the two interface designs (*game-style* vs. *industrial-style*) differ in learning performance?

Although the two designs were created in such a way that they differed only in aesthetic features that were not detrimental for learning (by considering multimedia design principles, see Mayer [37], Sweller [66] and usability, see Moshagen et al. [43]), we hypothesized, based on theories on aesthetic and emotional design [7,68,78], that the two interface designs would have a different impact on learners' affects and, thus, on (H4a) objective and (H4b) self-assessed learning performance. Moreover, we expected a general increase in (H4c) objective and (H4d) self-assessed knowledge after learning for both designs.

Research Question 5. (explorative) Does aesthetic experience (*surface structures* and *deeper perceptual processes*) have an impact on learning performance?

According to the Research Questions 1 and 4, and in line with related research on aesthetic art experience and user experience (see second research gap in section 1.3), we expected that aesthetic experience (i.e., "expressive aesthetics," "emotion," "cognitive stimulation") and usability would have an impact on (H5a) objective and (H5b) self-assessed learning performance.

To answer these research questions and hypotheses, we set up a field study (1) to investigate the effects of two different learning app designs on interest and learning performance and (2) to investigate the effects of three aesthetic experience dimensions on these dependent variables. In the next section, the method of the present study is described in detail.

2. Method

2.1. Participants and design

A total of 108 secondary school students (Switzerland) participated in the experiment as part of a pre-holiday school project during regular teaching hours. The ethical standards were set through the institutional ethics committee of our institution. The participants had no previous experience with the topic of study in the experiment at the time of the study and were randomly assigned to one of two conditions using different interface designs for a learning app: the *game-style* ($n = 53$) and the *industrial-style* design variant ($n = 55$), in the following referred to as *Interface Aesthetic* (see Fig. 1). The sample consisted of 53 female and 55 male teenagers ($M = 13.3$ years, $SD = 0.53$, range = 12 -15). The experimental design was mixed two-factorial, with *Interface Aesthetic* as between factor (interface design variants *game-style* vs. *industrial-style*) and *Time of Measurement* as within factor (pre-post-test). The primary dependent variables were interest (situational and individual) and learning performance (objective knowledge and subjective self-assessments). In addition, aesthetic experience was measured by a validated questionnaire to assess the impact of aesthetic experience (i.e., "expressive aesthetics," "emotion," "cognitive stimulation") on the dependent variables.

2.2. Materials

The learning material addressed the topic of *energy* (physics) and covered three subtopics that are typically part of the school curriculum for secondary school: energy forms (e.g., electrical energy), energy sources (e.g., sun, wind), and power plants (e.g., biomass power plant). The learning app included a geographic map showing a culture trail on industrial history – well known in the region were the study was conducted. On this trail, there is an option to visit historic buildings (e.g., a historic power plant). The app "virtually" guided students through three locations and offered context and background information about the spots, as well as further texts about energy forms, energy sources, and power plants. Additionally, short quizzes were integrated to foster a playful interaction with the learning app, which has been found to positively impact learning [e.g., 16]. Note, these quizzes were separate to the learning performance questionnaire (see section 2.3). The app was developed by a professional designer and was run on the iPad Air 2.

To create the design variants, a pilot study was conducted that followed several steps. First, a focus group was conducted, which included an expert discussion held with seven specialists from different interdisciplinary fields: contemporary design practices and aesthetic (art) design (two experts), user experience (one expert), interface- and interaction design (one expert), interactive media technologies (one expert), and instructional design (two experts). The goals of this workshop were to understand the concept of aesthetics based on the different expert opinions and the derivation of several criteria on the basis of which the interface variants were created. This involved criteria which had to be consistent in the design variants (i.e., usability and instructional design), in order to ensure optimal learning support and criteria, which were allowed to differ between the variants, i.e., aesthetic features not directly impacting learning through instructional advantages (referred to as *Interface Aesthetic*, see Table 1). Second, based on these criteria, a professional designer created four different interface designs that oriented towards different life worlds (see first study objective in section 1.4). Third, the four design variants were rated by the same group of experts with regard to the preassigned criteria: apart from the multimedia design principles and the criterion on custom design – for each of which it was necessary to indicate whether they were demonstrated in the design or not – the criteria were rated on a Likert scale from 1 to 7: from little to very (e.g., conceptual basis of text and figures in the design variants: little discernible to very well discernible) or on a scale with opposites (e.g., illustrations: very realistic to very abstract). Additionally, open responses of the experts to the criteria were considered qualitatively. Fourth, by means of descriptive analyses (due to the small number of participants) two interface designs were chosen for the present work that differed the least in their instructional design, their general quality, and their customization to the target group and the most in their *Interface Aesthetic* (i.e., aesthetic features that do not directly impact learning). For example, color was used in one of the design variants to increase appeal and attractiveness (see section 1.1), but did not directly support learning due to a signaling effect [cf. 37]. We contrasted a *game-style* design variant with an *industrial-style* variant (see Fig. 1). Note, outcomes from the focus group were also taken into account in the development of the aesthetic experience questionnaire (see section 2.4).

2.3. Measures

The number of items, ranges of Likert scale scores of all measurements, and their usage in pre- and post questionnaires can be seen in Table 2. Note, we describe our questionnaire for measuring aesthetic experience in a separate section (see section 2.4).

To measure *surface structures* with a validated (Cronbach's $\alpha = 0.76$) and well-known instrument, we used a short version of the Visual Aesthetics of Website Inventory [VisAWI-S, see 44].

Interest. In order to measure (1) situational and (2) individual interest (cf. section 1.2 for a definition), several scales were used: (1) situational interest was quantified using two measures that were identified as momentary experiences triggered by the learning environment: first, the sum of three absolute single-items measuring pre- and post-experimental interest to engage with each of the three learning topics [2]. Second, the intrinsic/enjoyment subscale of the short German version of the intrinsic motivation inventory (IMI) [77] was used to measure pre- and post-experimental interest to engage with the app (Cronbach's $\alpha = 0.89$).

To measure (2) individual interest, three measures were used that were characterized by a more long-lasting interest, probably leading to a persistent willingness to continue to study the learning material in the future: first, an absolute one-item scale measuring pre- and post-experimental general interest in physics, second, pre- and post-experimental interest in physics compared to other school subjects, and third, a post-experimental single item, where students could self-assess the change in their interest in physics after learning with the app [2].

Learning performance was measured with three types of tasks, according to Moreno and Mayer [42]: matching tasks, retention tasks, and transfer tasks. Matching tasks were available for each of the three learning topics (i.e., energy forms, energy sources, and power plants). The purpose of these tasks was to identify four out of 16 correct items related to the topic (e.g., the correct items for the topic power plants were pumped-storage power plant, generator, turbine, and biomass power plant). The matching tasks had a score between 0 and 16 points, depending on how many of the 16 items were correctly classified. Each correct item classification as "right" or "wrong" was counted. The internal consistencies of the scales (48 items used in the pre-test and post-test for measuring prior knowledge and matching tasks) were performed according to Everitt and Skrondal [17] and all items with a correlation value less than 0.3 were dropped. After exclusion of 12 low consistency items, the final test version included 36 items covering the three learning topics for the matching tasks. Cronbach's alphas for the 36 items were 0.97 (pre-test) and 0.93 (post-test). Retention and transfer tasks were conducted together with an expert in the field and consisted of six multiple choice questions with four answering options, each administered during post-test (3 questions with multiple correct answers and 3 questions with only one correct answer). The questions had a score range between 0 and 1 points; the questions with a single possible answer awarded either 1 (for correct) or 0 points (for incorrect answers), questions with several possible answers awarded 0, 0.5, or 1 points. Due to the small numbers of items, no measurement of the internal consistency was conducted.

Finally, *self-assessed knowledge* was measured by three single items about pre- and post-experimental self-assessed knowledge for each of the three learning topics (i.e., energy forms, energy sources, power plants) [2]. For further analyses, we summed the values of the three pre- and post-test variables to one pre- and one post-test variable. Hence, a total score of 18 points could be achieved in the pre- and post-test scores for self-assessed knowledge.

To test our hypotheses, we used these scales to (1) conduct (multivariate) analyses of variance and (2) multiple regression analyses. The conditions for the statistical tests were checked in advance.

2.4. Aesthetic experience questionnaire

In order to measure aesthetic experience with an instrument that considered both *surface structures* and *deeper perceptual processes*, we developed a questionnaire [61]. The questionnaire was previously validated with teenage students of our age of interest (N = 160, 43.13% female, age: $M = 13.7$, $SD = 1.04$) who learned with aesthetically appealing learning apps. The statistical validation (PAF, Promax) of the questionnaire (Cronbach's $\alpha = 0.86\text{-}0.91$) confirmed three dimensions of *aesthetic experience*: "expressive aesthetics" (Cronbach's $\alpha = 0.68\text{-}0.86$), "cognitive stimulation" (Cronbach's $\alpha = 0.83\text{-}0.86$), and "emotion" (Cronbach's $\alpha = 0.68\text{-}0.86$). Simultaneously, a short "usability" scale was validated within the same experiment (Cronbach's $\alpha = 0.74\text{-}0.84$). Separately from this, the questionnaire also involved the control item "I can solve the learning task with the learning app" which we used as a manipulation check (see section 3.1). The questionnaire was developed (1) based on relevant related research (e.g., "expressive aesthetics": [30]; "cognitive stimulation": [19], [25], [48]; "emotion": [30], [53]; and for usability see [40], [63]) and (2) in consideration of the results from the focus group described in section 2.2. Following earlier approaches by Leder et al. [31] and Marković [35], for the present study we assigned the dimension "expressive aesthetics" to *surface structures* and the dimensions "cognitive stimulation" and "emotion" to *deeper perceptual processes* to measure post-experimental aesthetic experience provoked by the interaction with one of the two design variants. Note, although the questionnaire considered measures of general user experience (cf. section 1.3), we assume that the term "aesthetic experience" is more appropriate for the newly developed questionnaire for two main reasons: (1) our focus was on the investigation of learners' experience resulting from their engagement and interaction with an "aesthetic" design (stimulus = appealing app interface) by means of a correspondingly validated instrument, and (2) the development of the design variants and the questionnaire was based on qualitative research methods, such as a focus group and expert group discussions, particularly focusing on measuring aesthetic experience (see also section 2.2).

2.5. Procedure

The study took place in June 2017 and was conducted on-site in class in a school in Switzerland and in German language. Overall, the experiment lasted about 90 minutes (two school lessons). The different phases of the procedure are summarized in more detail in Fig. 2. In the pre-experimental phase, students were welcomed and informed of the content of the study and received information according to the ethics standards approved by the ethics committee of our institution (15 min). Next, they were invited for pre-testing in their classrooms and worked individually on one desk. Students were guided through the pre-test questionnaire (paper-and-pencil), including demographics and prior interest and knowledge (10 min). Just before learning, students received a brief tutorial of the learning app and the task (10 min), before being randomly assigned to one of the two designs of the learning app (*Interface Aesthetic: game-style* vs. *industrial-style*). In the experimental phase, students had about 30 minutes to learn the three topics with the app (10 min for each learning topic). Although the time of the experiment was limited due to the embedding of the study directly in the classroom (within school schedule), students had enough time to complete learning. Finally, in the post-experimental phase that followed a 15-minute break, students were invited to answer the paper-and-pencil post-test questionnaires (20 min) before they were thanked and released.

3. Results

All data analyses were performed using SPSS.

3.1. Results of between-group comparisons and manipulation check

Chi-square test for gender and *t*-tests for age and prior media experience yielded no significant differences between conditions ($p > 0.10$).

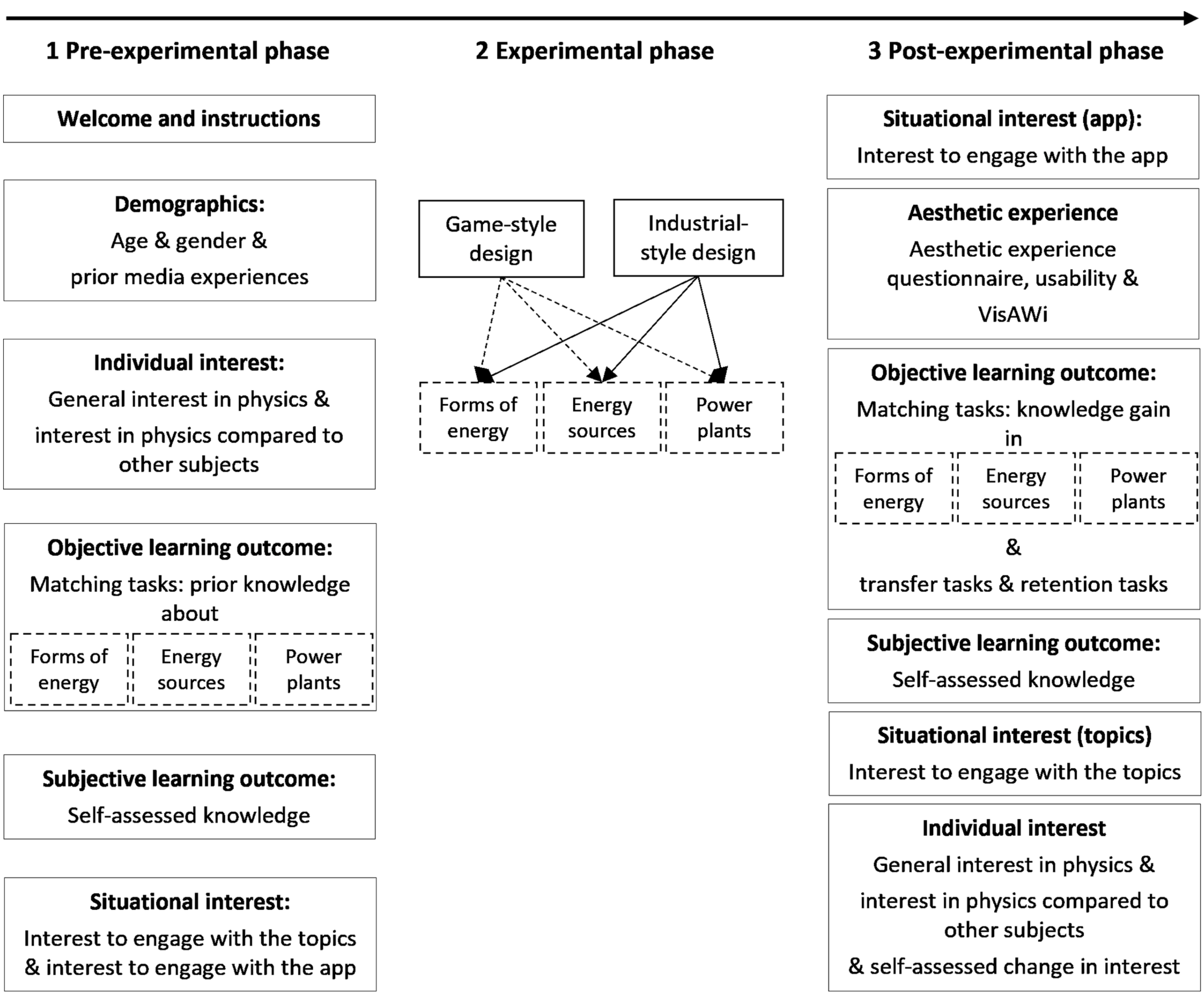

Fig. 2. The figure describes the procedure of the study.

Moreover, multivariate analysis of variance (MANOVA) with *Interface Aesthetic* (*game-style* vs. *industrial-style*) as between-subjects factor and prior knowledge and prior interest as dependent variables did not reach significance ($p > 0.10$). Hence, groups were comparable on these variables. In order to ensure the usability (see Table 2 for details) between the designs was equal and that we only varied *Interface Aesthetic* (*game-style* vs. *industrial-style*), a *t*-test for independent samples was conducted, revealing no significant difference between the design variants on usability, $t(106) = -0.06$, $p = 0.953$, $d = -0.01$. This indicated that students in both conditions experienced the app on about the same levels of usability (*game-style* variant: $M = 4.57$, $SD = 0.50$; *industrial-style* variant: $M = 4.57$, $SD = 0.49$). Moreover, results from a *t*-test for independent samples revealed no significant differences between the conditions regarding the control item "I can solve the learning task with the learning app", $t(105) = 0.80$, $p = 0.425$, $d = 0.16$. This indicates that students in both conditions found the app similarly suitable for completing the given task.

3.2. Interface aesthetic and aesthetic experience (RQ 1)

In order to investigate Research Question 1, we focused on the differences between the two designs (*Interface Aesthetic: game-style* vs. *industrial-style*) on aesthetic experience (see Table 3 for descriptive data).

To test Hypothesis 1a, whether learning with the *game-style* design would result in higher scores in *surface structures*, an independent sample *t*-test was conducted with mean scores of the VisAWI-S as dependent variable. The *t*-test revealed a significant result, $t(106) = -3.92$, $p < 0.01$, $d = 0.71$, indicating that (as expected) the students who learned with the *game-style* design rated the app significantly higher in its *surface structures* ($M = 5.45$, $SD = 1.02$) than students who learned with the *industrial-style* design ($M = 4.65$, $SD = 1.10$).

To test Hypothesis 1b, whether learning with the *game-style* design would result in higher scores of *aesthetic experience* (see Fig. 3), a MANOVA with the three dimensions of *aesthetic experience* (i.e., "expressive aesthetics," "emotion," "cognitive stimulation") as dependent variables was conducted and revealed no significant results ($F(3, 104) = 1.825$, $p = 0.147$; Wilks Lambda = 0.950). Yet, univariate post-hoc tests revealed a significant effect on "expressive aesthetics," $F(1, 106) = 3.998$, $p = 0.048$, $d = 0.38$, suggesting students who learned with the *game-style* design experienced higher "expressive aesthetics" ($M = 3.64$, $SD = 0.82$) than students who learned with the *industrial-style* design ($M = 3.31$, $SD = 0.90$).

3.3. Interface aesthetic and interest (RQ 2)

To investigate Research Question 2, whether the two interface designs (*game-style* vs. *industrial-style*) differ in their impact on students' situational and individual interest, several analyses were conducted (see Table 4 for descriptive data).

In order to investigate effects for situational interest (H2a, H2c), two mixed ANOVAs with *Interface Aesthetic* as between- and measuring time as within-subjects factor and either interest to engage with the topics or with the app as dependent variable were conducted. As expected (H2a), a significant effect was found for *Interface Aesthetic* regarding interest to engage with the topics ($F(1, 100) = 3.03$, $p = 0.0425$, Wilks Lambda = 0.971), indicating that students learning with the *game-style* design had a

Table 1
Criteria for designing the learning application .

Criteria committed to be similar (equal learning support)	Criteria permitted to deviate (Interface Aesthetic)
Instructional design	**Surface structures**
Principles of multimedia Learning	Color
Usability	Expressivity (classical to creative)
General application quality	Typography
General composition	Illustrations (realistic to abstract)
Craftsmanship	Gestalt
Conceptual basis	Orientation (playful to technical)
Text content and structure	
Target group	**Deeper perceptual processes**
Zeitgeisty	Expectable experience
Addressing preferences of target group	Cognitive stimulation
Custom Design	Emotional stimulation

Table 2
Index variables, measures, items and score of the used questionnaires.

Measure	Sample Items	Scores	Pre/ Post
Surface Structures			
VisAWI	4 items e.g.: "The layout is pleasantly varied"	1 (strongly disagree) - 7 (strongly agree)	post
Aesthetic Experience			
Surface structures and deeper processes			
Expressive aesthetics	3 items e.g.: "The special effects in the app are well done"	1 (strongly disagree) - 5 (strongly agree)	post
Cognitive stimulation	6 items e.g.: "I would love to explore the topic further with the app."		
Emotion	4 items e.g.: "Using the app is fun."		
Usability	3 items e.g.: "It is easy to navigate in the learning app."	1 (strongly disagree) - 5 (strongly agree)	post
Interest			
Situational interest			
Interest to engage with the learning topics	1 item for each learning topic: "It is important to me, to know what is behind X"	1 (strongly disagree) - 6 (strongly agree)	pre/ post
Interest to engage with the learning app	3 items e.g.: "I think I will enjoy/I enjoyed learning the topics with the app"	1 (strongly disagree) - 5 (strongly agree)	pre/ post
Individual interest			
General interest in physics	1 item: "To what extend are you interested in physics?"	1 (not at all) - 6 (very much)	pre/ post
Interest compared to other subjects	8 items e.g. "If you had to decide, what do you think, which of the school subjects do you like better?"	1 (subject 1) - 6 (subject 2)	pre/ post
Self-assessed change in interest	1 item: "In the first questionnaire we asked you about your interests and your opinion about physics. How has this opinion changed?"	-3 (rather worse) - 3 (rather improved)	post
Learning Performance			
Objective learning outcome			
Matching Tasks	16 items for each learning topic e.g.: "Please tick the terms which you would most likely assign to the topic XXX. Multiple answers may be checked."	16 terms per topic, of which 4 were respectively correct	pre/ post
Retention Tasks	4 items e.g.: "What type of power plant was the Kappelerhof power plant?"	4 answer options, 1 or multiple correct answers	post
Transfer Tasks	2 items e.g.: "A conventional electric oven... (answer:) converts the electrical energy into thermal energy."		post
Subjective learning outcome			
Self-assessed knowledge	1 item for each learning topic: "I know what X is/ are"	1 (strongly disagree) - 6 (strongly agree)	pre/ post

Table 3
Means and standard deviations for interface designs (*industrial-style* vs. *game-style*) for aesthetic experience.

	Industrial-style design	*Game-style* design
Surface Structures	M (SD)	M (SD)
VisAWI	4.65 (1.10)	5.45 (1.02)
Aesthetic Experience		
Cognitive Stimulation	3.12 (0.75)	3.18 (0.81)
Expressive Aesthetics	3.31 (0.90)	3.64 (0.82)
Emotion	3.95 (0.68)	3.93 (0.68)

Note: See Table 2 for score ranges.

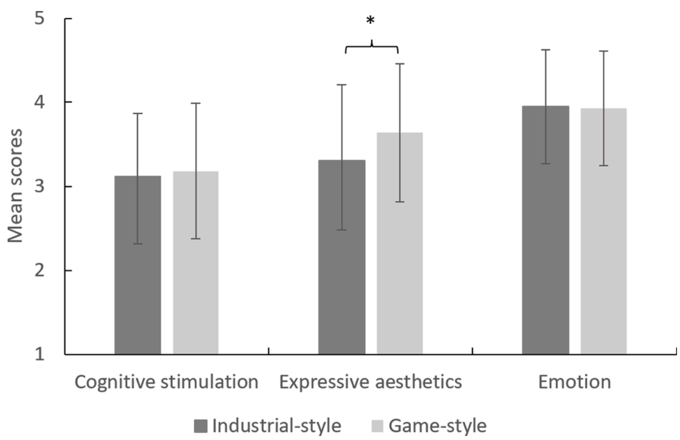

Fig. 3. Comparison between *game-style* and *industrial-style* on the three dimensions of aesthetic experience. * *game-style* design significantly assessed higher in "expressive aesthetics" than *industrial-style* design.

Table 4
Means and standard deviations for interface designs (*industrial-style* vs. *game-style*) for situational and individual interest measures.

	Industrial-style design		*Game-style* design	
	pre	post	pre	post
Situational interest	M (SD)	M (SD)	M (SD)	M (SD)
Interest to engage with the learning topics	3.47 (1.29)	3.59 (1.33)	3.29 (1.16)	3.76 (1.30)
Interest to engage with the learning app	3.55 (0.60)	3.80 (0.69)	3.44 (0.69)	3.63 (0.86)
Individual interest				
General interest in physics	3.71 (1.25)	3.83 (1.24)	3.91 (1.23)	3.98 (1.29)
Interest compared to other subjects	3.19 (0.87)	3.21 (0.89)	2.97 (0.78)	3.00 (0.79)
Self-assessed change in interest		0.66 (0.92)		1.02 (1.04)

Note: See Table 2 for score ranges.

higher increase in situational interest after learning than students learning with the *industrial-style* design. However, no between-subjects effect was found for interest to engage with the app ($p < 0.10$). Furthermore, according to expectations (H2c), significant effects for measuring time were found for both interest to engage with the topics ($F(1, 100) = 7.889$, $p < 0.01$, Wilks Lambda = 0.927) and interest to engage with the learning app ($F(1, 105) = 16.503$ $p < 0.01$, Wilks Lambda = 0.864), indicating an increase of situational interest after learning for both designs (see Table 4).

Regarding individual interest (H2b), a *t*-test for independent samples with the post single item self-assessed change in interest as dependent variable revealed a marginally significant effect, $t(103) = -1.877$, $p = 0.063$, demonstrating that students with the *game-style* design stated more positive change in individual interest after using the app than students who learned with the *industrial-style* design. However, mixed ANOVAs with *Interface Aesthetic* (*game-style* vs. *industrial-style*) as between- (H2b) and measuring time (H2d) as within-subjects variable and general interest in physics and interest in physics compared to other subjects as dependent variables did not reach statistical significance ($p < 0.10$).

3.4. Aesthetic experience and interest (RQ3)

To examine Research Question 3, if *aesthetic experience* has an impact on situational and individual interest, we took a closer look at the three dimensions "emotion," "cognitive stimulation," and "expressive aesthetics," and their relation with situational and individual interest. It is important to note that we additionally considered usability within the following analyses.

First, two multiple regression analyses with the three dimensions of aesthetic experience and usability as predictors, and the different values of the situational interest measures (H3a, interest to engage with the app and the topics) as dependent variables were computed. Both analyses revealed significant regression equations: interest to engage with the app: $F(4,102) = 4.341$, $p = 0.003$ with $R^2 = 0.145$, interest to engage with the topics: $F(4,46) = 3.451$, $p = 0.015$ with $R^2 = 0.231$. For both, the dimension "emotion" was found to predict significant change in interest (engage with the app: $B = 0.289$, $p = 0.007$; engage with the topics: $B = 2.091$, $p = 0.018$). These results indicate that students who experienced more positive emotions through design had a greater positive change in situational interest.

Similar multiple regression analyses with the dependent measures of individual interest (H3b) were then conducted. The analysis with students' self-assessed change in interest in physics as a dependent variable revealed a significant result ($F(4,100) = 5.692$, $p < 0.001$ with $R^2 = 0.182$): the more "cognitive stimulation" was experienced by the learners, the more positive was their self-assessed change in interest in physics ($B = 0.570$, $p < 0.001$). Multiple regression analyses with general interest in physics and interest in physics compared to other subjects did not reveal significant results ($p > 0.10$).

3.5. Interface aesthetic and learning performance (RQ4)

To investigate Research Question 4, regarding the effects of *Interface Aesthetic* (*game-style* vs. *industrial-style*) on objective (post-test: transfer, retention; pre-post-test: matching tasks) and self-assessed learning performance (pre-post-test), several tests were conducted.

To test hypotheses 4a and c, regarding objective learning performance, two analyses were run. First, a mixed 2x2 ANOVA with *Interface Aesthetic* as between- and measuring time as within-subjects factor with pre- and post-measurements of matching tasks was conducted. This showed a significant result for measuring time, $F(1,106) = 98.175$, $p < 0.001$, $d = 1.09$. According to expectations (H4c), this result indicates that students scored significantly higher in the post- ($M = 28.68$, $SD = 7.68$) than in the pretest ($M = 17.72$, $SD = 11.84$). No significant results were found for *Interface Aesthetic* ($p > 0.10$). Second, since transfer and retention tasks were highly correlated (Pearson: $B = 0.317$, $p = 0.001$), a MANOVA with *Interface Aesthetic* as a between-subjects factor was conducted. However, contrary to expectations (H4a), no main effects of *Interface Aesthetic* or any interaction effects concerning these measures yielded significance ($p > 0.10$).

In order to test hypotheses 4b and d, dealing with self-assessed knowledge, a mixed 2x2 ANOVA with self-assessed knowledge as dependent and *Interface Aesthetic* as between- and measuring time as within-subjects factor was conducted. The analysis yielded a highly significant effect for measuring time, $F(1, 106) = 215.819$, $p < 0.001$ Wilks Lambda = 0.329, indicating an increase in self-assessed knowledge (as expected H4c). Contrary to our expectations (H4d), no effect for *Interface Aesthetic* was found ($p > 0.10$).

3.6. Aesthetic experience and learning performance (RQ5)

In order to investigate Research Question 5, which regards the effects of aesthetic experience on learning performance, we inspected the effects of the dimensions "expressive aesthetics," "emotion," and "cognitive stimulation," as well as usability on objective and self-assessed learning performance. To determine the effects on objective learning performance (H5a), several multiple regression analyses with transfer, retention, and the difference value of matching tasks as dependent variables were conducted. The analysis computed on transfer tasks revealed a significant regression equation, $F(4,94) = 2.662$, $p = 0.037$, with $R^2 = 0.102$. However, while the dimensions of aesthetic experience showed no significant result, usability significantly predicted outcome in transfer tasks ($B = 0.107$, $p = 0.003$). This result indicates that the higher the usability of the app, the better transfer task outcomes. Moreover, the regression equation of matching tasks showed a significant result, $F(4,94) = 2.618$, $p = .040$ with $R^2 = 0.100$. It was found that "cognitive stimulation" significantly predicted outcome in matching tasks ($B = -.897$, $p = 0.011$). Interestingly, this result indicates that the *less* "cognitive stimulation" was perceived, the better the students' knowledge gains in matching tasks. Finally, the analysis conducted on retention tasks did not reach significance ($p > 0.10$).

To investigate effects of the aesthetic experience dimensions and usability on self-assessed knowledge (H5b), a multiple regression analysis was conducted with subjective knowledge after learning with the app (post-test) as the dependent variable. The analysis revealed a significant regression equation, $F(4,92) = 7.613$, $p < 0.001$ with $R^2 = 0.237$. It was found that "cognitive stimulation" ($B = 0.349$, $p < 0.001$) and usability ($B = 0.546$, $p = 0.010$) significantly predicted self-assessed knowledge: the more cognitive stimulating and the higher the usability of the app was perceived, the higher were ratings in self-assessed knowledge.

3.7. Interim summary of results

Before the results are discussed in detail, they are briefly summarized in the following. First, our results indicate that learning with both designs led to higher situational interest and higher self-assessed and objective learning performance immediately after learning.

Second, regarding the effects of appealing and learning-supportive interface designs (first study objective, cf. section 1.4), we found that *Interface Aesthetic* had an impact on perceived *surface structures* (VisAWI and dimension "expressive aesthetics" of aesthetic experience) in such a way that the *game-style* design was rated higher than the *industrial-style* design. Learning with the *game-style* design further led to a higher increase in situational and individual interest (marginal) compared to learning with the *industrial-style* design. However, *Interface Aesthetic* had no direct impact on self-assessed or objective learning performance.

Third, regarding the effects of students' aesthetic experience (second study objective, cf. section 1.4), our results suggest that higher experience of "emotion" led to higher situational interest, while higher "cognitive stimulation" led to higher individual interest. Finally, concerning learning performance, our findings show that effects of "cognitive stimulation" were conflicting; while higher "cognitive stimulation" seems to improve self-assessed knowledge, higher results in objective learning performance (matching tasks) are related to *less* "cognitive stimulation." Moreover, a positive relation of perceived usability with higher scores in self-assessed and objective learning performance (transfer tasks) was found.

4. Discussion

By pursing the three following objectives, we intended to shed light onto the effects of aesthetic designs on interest and learning in science and the role of aesthetic experience. We aimed to systematically and holistically investigate the aesthetic construct by comparing two aesthetic interface designs of a learning application that both optimally supported learning (*Interface Aesthetic: game-style* vs. *industrial-style*). We also aimed to investigate (explorative) the impact of aesthetic experience on interest and learning by considering *surface structures* and *deeper perceptual processes* with the three dimensions "expressive aesthetics," "emotion," and "cognitive stimulation." Additionally, by focusing on teenagers, we intended to contribute new results on the effects of aesthetic designs in young participants with whom science literacy can be promoted particularly well [73] and where students' orientation regarding science topics can successfully be encouraged [1,58].

To achieve these objectives, a controlled field study in the classroom context was conducted. We examined the effects of two interface design variants on interest and learning and investigated relations between three dimensions of *aesthetic experience* (i.e., *surface structures*: "expressive aesthetics;" *deeper processes*: "emotion," "cognitive stimulation") and these dependent variables. The study had its limitations, which we will outline in detail below (cf. section 4.4). Nevertheless, our data revealed interesting results that will be discussed in the following sections.

4.1. Aesthetic experience of interface designs

Our data demonstrates that students learning with the *game-style* design rated the design higher in *surface structures* (VisAWI and dimension "expressive aesthetics" of aesthetic experience) than students learning with the *industrial-style* design. This result is in line with previous research indicating that aesthetic features such as color [e.g., 52] and expressivity [30] impact perceived attractiveness of an object (see section 1.1). However, no differences between the designs concerning *deeper perceptual processes* of aesthetic experience such as "emotion" and "cognitive stimulation" could be found. A possible explanation for this result is the following. Our definition of *aesthetic experience* was mainly based on approaches relating to the experience of artwork [31,35]. In contrast, our object of investigation was the interface design of learning applications. In fact, aesthetics and the resulting experiences, such as "emotion" and "cognitive stimulation" play a more important role in museums than in traditional learning settings, since such experiences are the main goal of art enthusiasts and museum visitors. On the contrary, when it comes to learning applications, *deeper processes* related to an aesthetic object may only play a minor role. However, in the next sections, it becomes clear that *deeper processes* are important when more goal-oriented constructs such as interest and learning performance are involved.

4.2. Impact of interface aesthetic and aesthetic experience on interest

Our results show that *Interface Aesthetic* has a direct effect on situational interest in such a way that the *game-style* design leads to a higher increase in situational interest than the *industrial-style* design. This finding indicates that interface designs that lead to positive experiences of *surface structures* (VisAWI and "expressive aesthetic") of the learning application (higher in the *game-style* design) influence the momentary experience that is reflected in teenagers' situational interest [20]. This result is also in line with research on emotional design [7,78]. Moreover, our results suggest that situational interest is related to the *deep perceptual process* "emotion" – the higher the emotional experience, the higher the situational interest to engage with the learning topics and the app. This is in line with Hidi and Renninger [24], indicating that the higher the affect elicited by the experience, the higher the perceived interest.

In contrast, we could not confirm that one of the two interface designs led to higher individual interest (direct effect of *Interface Aesthetic*). While self-assessed change in interest was descriptively higher in the *game-style* than in the *industrial-style* design condition, students were not more interested in the learning topic in general. However, the *deep* *perceptual process* ”cognitive stimulation” was positively related to individual interest. Individual interest – characterized as a persistent willingness to return to a certain object (such as the learning topic and application) over time [see 20] – is thus mainly dependent on the stimulation, involvement, and excitement of the aesthetic learning environment. These results further highlight the importance of considering aesthetic experience – especially *deeper processes* – as they provide evidence that a successful processing of aesthetics according to Leder et al. [31] is able to increase interest.

Furthermore, contrary to our results regarding situational interest, individual interest did not increase after learning. This is, however, not surprising, since it takes a long time for individual interest to develop and change and those changes are not measurable directly after the experiment. Further and reiterated experiments could address this point (see also section 4.4).

But how do the different results between situational and individual interest come about? According to Ekman [14], emotions are characterized by a short duration – as is situational interest [20]. It is therefore not surprising that these concepts are related. Moreover, our results indicate that *surface structure* features (such as color) of interface designs also tend to be related to such short-lasting states. In contrast, interface design and ”emotion” did not have a direct impact on long-lasting individual interest. It can be argued that stimulating and involving interfaces that lead to ”cognitive stimulation” have a more long-term impact on interest (i.e., individual interest), while designs that evoked more positive experiences of *surface structures* and lead to higher emotional experience have a more short-term impact on interest (i.e., situational interest).

4.3. Impact of interface aesthetic and aesthetic experience on learning performance

Research on emotional design [78] suggests that specific aesthetic related aspects are responsible for aesthetics, positively affecting motivation and interest and subsequently learning. According to this research, we could confirm with our study that aesthetics has an impact on interest (more precisely on situational interest, cf. section 3.3), but not on objective or subjective learning performance. One explanation for this result could be that we manipulated several aesthetic features not necessarily defined as emotional designs (cf. Table 1). Another possible reason could be found in the design standards of our interfaces: the designs – both high-quality designs – were constructed so that they differed only in aesthetic features not detrimental for learning [successful manipulation check and equal consideration of multimedia design principles, see 37,66] and therefore offered an environment that equally supported learning in both designs. This is supported, on one hand, by our results demonstrating that higher perceived usability predicted transfer knowledge and self-assessed knowledge scores (see section 3.6), and, on the other hand, by our results showing that subjective and objective learning performance increased after learning with both interface designs. The two designs differed in *surface structures* (VisAWI and ”expressive aesthetics”), but did not trigger *deeper perceptual processes* in learners (such as ”emotion” and ”cognitive stimulation”). Hence, we assume two related explanations: (1) interface designs that are not detrimental for learning are experienced differently, but the experience is confined to objective *surface structures* (such as color, cf. section 1.1), which are known to have a positive impact on appeal and attractiveness, and in turn, possibly on interest (cf. sections 1.1 and 1.2); (2) such designs do not have a direct impact on *deeper processes* (such as ”emotion” and ”cognitive stimulation”) that may be related to learning performance (cf. section 3.6).

A further reason for this result can be found in the fact that the *industrial-style* design more closely resembled to a classical school book and the *game-style* design was more inspired by gaming apps. Therefore, the *industrial-style* variant probably triggered the learning context associations in learners more, whereas the *game-style* design triggered a leisure context. Learning could, therefore, be facilitated in the *industrial-style* design, since the design was a familiar stimuli in this context, whereas the *game-style* design was a non-familiar stimuli and more likely led to an orienting response compared to the familiar stimuli [70]. Positive motivational factors induced by the *game-style* design that possibly had influenced learning performance could thereby be diminished. This assumption is also in keeping with our results showing that less stimulation, involving, and excitement (i.e., ”cognitive stimulation”) induced by the learning environment led to higher results in objective learning performance. Students perceiving a higher ”cognitive stimulation” induced by the learning environment are probably more occupied with processing the interface than the learning material. Important to note here is also the so-called seductive details effect, suggesting a negative impact of appealing but irrelevant information on learning [38]. However, on the contrary, students who perceived high ”cognitive stimulation” rated their self-assessed knowledge higher. This result may be related with our findings concerning the positive relation of ”cognitive stimulation” and individual interest, since self-assessed knowledge may be more related to stimulation and excitement – similar to interest – than to objective knowledge.

4.4. Limitations and future work

To investigate our study goals, we conducted a field study directly in the classroom – the familiar learning environment of the students. However, this approach is also accompanied by some limitations. Firstly, possible confounders applying in the field could not be excluded (e.g., students sharing learning information in the break, see section 2.4). Secondly, for organisational and resource reasons, the study was carried out in a single school and thus had a limited number of students participating. Therefore, future research should consider either involving multiple schools in the investigations or replicating the study in a more controlled laboratory environment. Furthermore, we intended to holistically investigate the effect of the construct aesthetic by the systematic investigation of two high-quality design variants that were both designed to optimally support learning. The designs were based on different themes and metaphors that reflected different life worlds of the students: a variant orientating towards mobile games (i.e., *game-style* variant) and one that represented the look and feel of a classical school book (i.e., *industrial-style* variant). Learning might, therefore, have been influenced by learners’ personal experiences with similar designs. For example, game-experienced learners might have benefited more (or less) from the *game-style* design. It might be interesting to consider such student characteristics in future research. Moreover, through the holistic approach, an investigation of the overall expression of a design and its influence on interest and learning is difficult to grasp, due to the complexity of all influencing factors. The extent to which approaches from Gestalt theory can provide answers has been investigated in related research [3,60] and should be addressed in more detail in future research. Additionally, when simultaneously manipulating various aesthetic features, it is difficult to trace back effects of single features. For example, it remains unclear whether the different amount of contrast between the variants resulting from color variations (i.e., isoluminant) possibly led to interaction/bias effects between aesthetics and usability. Moreover, the design of the interfaces also affected the displayed information graphics in the apps (e.g., representation of electrical energy, see Fig. 1). Therefore, it cannot be completely excluded that the graphics of one design variant could be better processed by the learners than the graphics of the other design. Furthermore, the present study focused only on short-term measures of interest and learning performance. However, both concepts – especially individual interest – develop and change over time. It is, therefore, of utmost importance to additionally involve long-term measures in future research. Our results on individual and situational interest should further be interpreted with caution as we mainly used not-validated single-item scales. Future studies should use standardized and validated scales specifically created to measure individual and situational interest. Moreover, although our study gives first evidence underlining the importance of considering aesthetic experience when investigating effects of aesthetic design on learning and interest, future research should increasingly include such measurements to further investigate the concept and its relevance for this research field. Also, physics, the chosen science topic in the present study, is very multifaceted. Thus, future work should consider conducting additional experiments involving formulas or more abstract concepts to investigate which aesthetic characteristics of interfaces better convoy and make the understanding of the topic easier. Finally, there is a major gap between female and male learners in science topics [e.g., 75], which was not considered in this study. Research should address this gender gap by investigating possible different interest-fostering aspects between female and male learners in aesthetic design and also science topics themselves, while being mindful to consider educational and cultural biases.

5. Conclusion

There is still a need to foster teenagers' interest in science. This can not only foster scientific literacy to increase students' awareness and willingness to engage in globally important issues, but can also encourage students to pursue a future career in the sciences. With the present field study, we aimed to find ways to foster teenage students' interest and learning of science with a systematic investigation of the impact of two appealing and learning-supportive aesthetic interface designs of a learning application. Moreover, we intended to provide new insights into the construct aesthetic experience by examining experiences of *surface structures* (i.e., color, complexity) and *deeper perceptual processes* (i.e., "emotion" and "cognitive stimulation") and investigate impact on interest and learning. Our study revealed that interface aesthetics can influence interest and learning in differently designed learning app variants, even when these variants provide an equally optimal learning support. More precisely, our results suggest that designs with high rated *surface structures* can positively impact interest to engage with the learning topic. Thus, important implications for the design of learning apps focusing on science can be derived. Moreover, we found that *deeper perceptual processes* – such as "emotion" and "cognitive stimulation" induced by the aesthetics of the design – are related to interest, as well as to self-assessed and objective learning performance. This highlights the importance of considering aesthetic experience, particularly such deeper processes, in future research.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

This publication was funded by the University of Applied Sciences and Arts Northwestern Switzerland (FHNW) within the Strategic Initiative "EduNaT" (2015 - 2017). The authors thank the University of Applied Science and Arts for its financial support. Moreover, the authors would like to thank Ulrike Schock and Matej Zmitko for designing the interfaces and Moritz Dietsch for programming the learning application. A special thanks goes also to our project partners Prof. Dr. Nicolaj vander Meulen and Prof. Dr. Jörg Wiesel. Last, we would like to thank Anna Caltabiano for proofreading.

References

[1] Almeda MVQ, Baker RS. Predicting student participation in STEM careers: the role of affect and engagement during middle school. Journal of Educational Data Mining 2020;12(2):33–47. https://doi.org/10.5281/zenodo.4008054.

[2] Tanner, Alexandra, Vollenwyder, B., Schock, U., Kalt, M., Mateescu, M., Agotai, D., Gros, P., Vogel, M., & Zahn, C. G. (2014). Interact! An Evidence-Based Framework For Digitally Supported School Field Trips. *Proceedings of the Ninth ACM International Conference on Interactive Tabletops and Surfaces*, 357–362. https://doi.org/10.1145/2669485.2669546.

[3] Bae J, Watson B. Toward a better understanding and application of the principles of visual communication. Handbook of human centric visualization. Springer; 2014. p. 179–201.

[4] Bencze L, Carter L. Globalizing students acting for the common good. J Res Sci Teach 2011;48(6):648–69. https://doi.org/10.1002/tea.20419.

[5] Bernacki ML, Nokes-Malach TJ, Aleven V. Examining self-efficacy during learning: variability and relations to behavior, performance, and learning. Metacogn Learn 2015;10(1):99–117. https://doi.org/10.1007/s11409-014-9127-x.

[6] Blijlevens J, Thurgood C, Hekkert P, Chen L-L, Leder H, Whitfield TWA. The aesthetic pleasure in design scale: the development of a scale to measure aesthetic pleasure for designed artifacts. Psychol Aesthet Creat Arts 2017;11(1):86–98. https://doi.org/10.1037/aca0000098.

[7] Brom C, Stárková T, D'Mello SK. How effective is emotional design? a meta-analysis on facial anthropomorphisms and pleasant colors during multimedia learning. Educational Research Review 2018;25:100–19. https://doi.org/10.1016/j.edurev.2018.09.004.

[8] Brooke J. Sus: a quick and dirty'usabilityvol. 189; 1996.

[9] Celik P, Onder F, Silay I. The effects of problem-based learning on the students' success in physics course. Procedia - Social and Behavioral Sciences 2011;28:656–60. https://doi.org/10.1016/j.sbspro.2011.11.124.World Conference on Educational Technology Researches - 2011

[10] Chen J, Wang M, Kirschner PA, Tsai C-C. The role of collaboration, computer use, learning environments, and supporting strategies in CSCL: a meta-analysis. Rev Educ Res 2018;88(6):799–843. https://doi.org/10.3102/0034654318791584.

[11] Desmet PMA, Hekkert P. Framework of product experience. International journal of design 2007;1(1):57–66.

[12] Dewey J. Art as experience. New York: Putnam's Sons; 1958.

[13] Drummond C, Fischhoff B. Individuals with greater science literacy and education have more polarized beliefs on controversial science topics. Proceedings of the National Academy of Sciences 2017;114(36):9587–92. https://doi.org/10.1073/pnas.1704882114.

[14] Ekman P. An argument for basic emotions. Cognition & emotion 1992;6(3–4):169–200.

[15] Etkina E, Van Heuvelen A, White-Brahmia S, Brookes DT, Gentile M, Murthy S, Rosengrant D, Warren A. Scientific abilities and their assessment. Physical Review special topics-physics education research 2006;2:020103. https://doi.org/10.1103/PhysRevSTPER.2.020103.

[16] Evans C, Gibbons NJ. The interactivity effect in multimedia learning. Computers & Education 2007;49(4):1147–60. https://doi.org/10.1016/j.compedu.2006.01.008.

[17] Everitt B, Skrondal A. The Cambridge dictionary of statisticsvol. 106. Cambridge University Press Cambridge; 2002.

[18] Feder T. Convincing US states to require physics. Phys Today 2011;64(7):29. https://doi.org/10.1063/PT.3.1163.

[19] Hager M, Hagemann D, Danner D, Schankin A. Assessing aesthetic appreciation of visual artworks—the construction of the art reception survey (ARS). Psychol Aesthet Creat Arts 2012;6(4):320–33. https://doi.org/10.1037/a0028776.

[20] Harackiewicz JM, Hulleman CS. The importance of interest: the role of achievement goals and task values in promoting the development of interest. Soc Personal Psychol Compass 2010;4(1):42–52. https://doi.org/10.1111/j.1751-9004.2009.00207.x.

[21] He L, Chen Y, Xiong X, Zou X, Lai K. Does science literacy guarantee resistance to health rumors? the moderating effect of self-efficacy of science literacy in the relationship between science literacy and rumor belief. Int J Environ Res Public Health 2021;18(5). https://doi.org/10.3390/ijerph18052243.

[22] Hekkert P, Leder H. Product aesthetics. In: Schifferstein HNJ, Hekkert P, editors. Product Experience. Amsterdam: Elsevier Science; 2007. p. 259–86.

[23] Hidi S, Anderson V. Situational interest and its impact on reading and expository writing. The role of interest in learning and development 1992;11:213–4.

[24] Hidi S, Renninger KA. The four-phase model of interest development. Educ Psychol 2006;41(2):111–27. https://doi.org/10.1207/s15326985ep4102_4.

[25] Isen AM, Reeve J. The influence of positive affect on intrinsic and extrinsic motivation: facilitating enjoyment of play, responsible work behavior, and self-control. Motiv Emot 2005;29(4):295–323. https://doi.org/10.1007/s11031-006-9019-8.

[26] Javora O, Hannemann T, Stárková T, Volná K, Brom C. Children like it more but don't learn more: effects of esthetic visual design in educational games. British Journal of Educational Technology 2019;50(4):1942–60. https://doi.org/10.1111/bjet.12701.

[27] Jung YJ, Zimmerman HT, Land SM. Emerging and developing situational interest during children's tablet-mediated biology learning activities at a nature center. Sci Educ 2019;103(4):900–22. https://doi.org/10.1002/sce.21514.

[28] Kocabalil AB, Laranjo L, Coiera E. Measuring user experience in conversational interfaces: a comparison of six questionnaires. Proceedings of the 32nd International BCS Human Computer Interaction Conference 32. 2018. p. 1–12.

[29] Krapp A. Structural and dynamic aspects of interest development: theoretical considerations from an ontogenetic perspective. Learn Instr 2002;12(4):383–409. https://doi.org/10.1016/S0959-4752(01)00011-1.

[30] Lavie T, Tractinsky N. Assessing dimensions of perceived visual aesthetics of web sites. Int J Hum Comput Stud 2004;60(3):269–98. https://doi.org/10.1016/j.ijhcs.2003.09.002.

[31] Leder H, Belke B, Oeberst A, Augustin D. A model of aesthetic appreciation and aesthetic judgments. British Journal of Psychology 2004;95(4):489–508. https://doi.org/10.1348/0007126042369811.

[32] Lindgaard G, Fernandes G, Dudek C, Brown J. Attention web designers: you have 50 milliseconds to make a good first impression! Behaviour & Information Technology 2006;25(2):115–26. https://doi.org/10.1080/01449290500330448.

[33] Linnenbrink-Garcia L, Durik AM, Conley AM, Barron KE, Tauer JM, Karabenick SA, et al. Measuring situational interest in academic domains. Educ Psychol Meas 2010;70(4):647–71. https://doi.org/10.1177/0013164409355699.

[34] Liu Y. Engineering aesthetics and aesthetic ergonomics: theoretical foundations and a dual-process research methodology. Ergonomics 2003;46(13–14):1273–92. https://doi.org/10.1080/00140130310001610829.

[35] Markovi S. Components of aesthetic experience: aesthetic fascination, aesthetic appraisal, and aesthetic emotion. Iperception 2012;3(1):1–17. https://doi.org/10.1068/i0450aap.

[36] Mautone PD, Mayer RE. Signaling as a cognitive guide in multimedia learning. J Educ Psychol 2001;93(2):377389. https://doi.org/10.1037/0022-0663.93.2.377.

[37] Mayer RE. Cognitive theory of multimedia learning. 2. Cambridge: Cambridge University Press; 2014, ISBN 9781107610316. p. 43–71.

[38] Mayer RE. Taking a new look at seductive details. Appl Cogn Psychol 2019;33(1): 139–41. https://doi.org/10.1002/acp.3503.

[39] Mayer RE, Estrella G. Benefits of emotional design in multimedia instruction. Learn Instr 2014;33:12–8. https://doi.org/10.1016/j.learninstruc.2014.02.004.

[40] Minge M, Riedel L, Thüring M. Modulare evaluation von technik. entwicklung und validierung des meCUE fragebogens zur messung der user experience [modular evaluation of technology. development and validation of the meCUE questionnaire for measuring user experience]. In: Brandenburg E, Doria L, Gross A, Güntzler T, Smieszek H, editors. Grundlagen und Anwendungen der Mensch-Technik-Interaktion. Berlin: University Press of TU Berlin; 2013. p. 28–36.

[41] Moreno R. Does the modality principle hold for different media? a test of the method-affects-learning hypothesis. Journal of Computer Assisted Learning 2006; 22(3):149–58. https://doi.org/10.1111/j.1365-2729.2006.00170.x.

[42] Moreno R, Mayer RE. Cognitive principles of multimedia learning: the role of modality and contiguity. J Educ Psychol 1999;91(2):358. https://doi.org/10.1037/0022-0663.91.2.358.

[43] Moshagen M, Musch J, Göritz AS. A blessing, not a curse: experimental evidence for beneficial effects of visual aesthetics on performance. Ergonomics 2009;52(10): 1311–20. https://doi.org/10.1080/00140130903061717.

[44] Moshagen M, Thielsch M. A short version of the visual aesthetics of websites inventory. Behaviour & Information Technology 2013;32(12):1305–11. https://doi.org/10.1080/0144929X.2012.694910.

[45] Moshagen M, Thielsch MT. Facets of visual aesthetics. Int J Hum Comput Stud 2010;68(10):689–709. https://doi.org/10.1016/j.ijhcs.2010.05.006.

[46] Norman DA. Emotion and design: attractive things work better. Interactions Magazine 2002;9(4):36–42.

[47] Norman DA. Emotional design: why we love (or hate) everyday things. New York: Basic Books; 2004. https://doi.org/10.1145/966012.966013.

[48] O'Brien HL, Toms EG. The development and evaluation of a survey to measure user engagement. Journal of the American Society for Information Science and Technology 2010;61(1):50–69. https://doi.org/10.1002/asi.21229.

[49] OECD. Assessing scientific, reading and mathematical literacy. 2006. https://doi.org/10.1787/9789264026407-en.

[50] OECD. Education at a glance 2018. 2018. https://doi.org/10.1787/eag-2018-en.

[51] Oliver MC, Adkins MJ. æHot-headedǥ students? scientific literacy, perceptions and awareness of climate change in 15-year olds across 54 countries. Energy Research & Social Science 2020;70:101641. https://doi.org/10.1016/j.erss.2020.101641.

[52] Palmer SE, Schloss KB, Sammartino J. Visual aesthetics and human preference. Annu Rev Psychol 2013;64(1):77–107. https://doi.org/10.1146/annurev-psych-120710-100504.

[53] Pekrun R, Goetz T, Perry RP. Achievement emotions questionnaire (AEQ). user's manual. Unpublished Manuscript, University of Munich, Munich 2005.

[54] Plass JL, Heidig S, Hayward EO, Homer BD, Um E. Emotional design in multimedia learning: effects of shape and color on affect and learning. Learn Instr 2014;29: 128–40. https://doi.org/10.1016/j.learninstruc.2013.02.006.

[55] Plass JL, Kalyuga S. Four ways of considering emotion in cognitive load theory. Educ Psychol Rev 2019;31(2):339–59. https://doi.org/10.1007/s10648-019-09473-5.

[56] Plass JL, Kaplan U. Chapter 7 - emotional design in digital media for learning. In: Tettegah SY, Gartmeier M, editors. Emotions, technology, design, and learning. San Diego: Academic Press; 2016. p. 131–61. https://doi.org/10.1016/B978-0-12-801856-9.00007-4.

[57] Reber R, Schwarz N, Winkielman P. Processing fluency and aesthetic pleasure: is beauty in the perceiver's processing experience? Personality and Social Psychology Review 2004;8(4):364–82. https://doi.org/10.1207/s15327957pspr0804_3.

[58] Reinhold S, Holzberger D, Seidel T. Encouraging a career in science: a research review of secondary schools' effects on students' STEM orientation. Studies in Science Education 2018;54(1):69–103. https://doi.org/10.1080/03057267.2018.1442900.

[59] Renninger KA, Hidi S. Interest, attention, and curiosity. In: Renninger KA, Hidi S, editors. The power of interest for motivation and engagement. New York, NY: Routledge; 2016. p. 32–51.

[60] Reynoso JMG, Romo LIS. Measuring the effectiveness of designing end-user interfaces using design theories. International Journal of Information Technologies and Systems Approach (IJITSA) 2020;13(2):54–72. https://doi.org/10.4018/IJITSA.2020070103.

[61] Roethlisberger B.. Interface Ästhetik und design von digitalen lernmedien: Konstruktion eines fragebogens zur erfassung des ästhetischen empfindens [interface aesthetics and design of digital learning media: Construction of a questionnaire for the assessment of aesthetic experience]; 2017. Unpublished master's thesis.

[62] Romine W, Tsai C-L, Miller M, Tang N-E, Folk W. Evaluation of a process by which individual interest supports learning within a formal middle school classroom context. International Journal of Science and Mathematics Education 2020;18(7): 1419–39. https://doi.org/10.1007/s10763-019-10032-1.

[63] Sauro J. Supr-q: a comprehensive measure of the quality of the website user experience. J Usability Stud 2015;10(2):68–86.

[64] Schneider S, Häßler A, Habermeyer T, Beege M, Rey GD. The more human, the higher the performance? examining the effects of anthropomorphism on learning with media. J Educ Psychol 2019;111(1):57. https://doi.org/10.1037/edu0000273.

[65] Seckler M, Opwis K, Tuch AN. Linking objective design factors with subjective aesthetics: an experimental study on how structure and color of websites affect the facets of users' visual aesthetic perception. Comput Human Behav 2015;49: 375–89. https://doi.org/10.1016/j.chb.2015.02.056.

[66] Sweller J. Implications of cognitive load theory for multimedia learning. The Cambridge handbook of multimedia learning 2005;3(2):19–30.

[67] Thielsch MT, Niesenhaus J. User experience, gamification, and performance. The Wiley Blackwell handbook of the psychology of the internet at work 2017:79–102.

[68] Thielsch MT, Scharfen J, Masoudi E, Reuter M. Visual aesthetics and performance: A first meta-analysis. Proceedings of Mensch und Computer 2019. 2019. p. 199–210. https://doi.org/10.1145/3340764.3340794.

[69] Tuch AN, Presslaber EE, Stöcklin M, Opwis K, Bargas-Avila JA. The role of visual complexity and prototypicality regarding first impression of websites: working towards understanding aesthetic judgments. Int J Hum Comput Stud 2012;70(11): 794–811. https://doi.org/10.1016/j.ijhcs.2012.06.003.

[70] Uhl F, Lang W, Spieth F, Deecke L. Negative cortical potentials when classifying familiar and unfamiliar faces. Cortex 1990;26(1):157–61. https://doi.org/10.1016/S0010-9452(13)80082-2.

[71] Uitto A. Interest, attitudes and self-efficacy beliefs explaining upper-secondary school students' orientation towards biology-related careers. International Journal of Science and Mathematics Education 2014;12(6):1425–44. https://doi.org/10.1007/s10763-014-9516-2.

[72] Um E, Plass JL, Hayward EO, Homer BD, et al. Emotional design in multimedia learning. J Educ Psychol 2012;104(2):485. https://doi.org/10.1037/a0026609.

[73] Vieira RM, Tenreiro-Vieira C. Fostering scientific literacy and critical thinking in elementary science education. International Journal of Science and Mathematics Education 2016;14(4):659–80. https://doi.org/10.1007/s10763-014-9605-2.

[74] Wang J, Hsu Y. The relationship of symmetry, complexity, and shape in mobile interface aesthetics, from an emotional perspectivea case study of the smartwatch. Symmetry (Basel) 2020;12(9). https://doi.org/10.3390/sym12091403.

[75] Wang M-T, Degol JL. Gender gap in science, technology, engineering, and mathematics (STEM): current knowledge, implications for practice, policy, and future directions. Educ Psychol Rev 2017;29(1):119–40. https://doi.org/10.1007/s10648-015-9355-x.

[76] Wang Y, Zhou Z, Gong S, Jia D, Lei J. The effects of emotional design on multimedia learning and appreciation of chinese poetry. Front Psychol 2021;12. https://doi.org/10.3389/fpsyg.2021.621969.621969–621969

[77] Wilde M, Bätz K, Kovaleva A, Urhahne D. Testing a short scale of intrinsic motivation. Zeitschrift für Didaktik der Naturwissenschaften 2009;15.

[78] Wong RM, Adesope OO. Meta-analysis of emotional designs in multimedia learning: a replication and extension study. Educ Psychol Rev 2021. https://doi.org/10.1007/s10648-020-09545-x.

[79] Zimmerman BJ. Becoming a self-regulated learner: an overview. Theory Pract 2002;41(2):64–70.