Education and Information Technologies (2024) 29:16051–16072
https://doi.org/10.1007/s10639-024-12460-1

The influence of content-relevant background color as a retrieval cue on learning with multimedia

Felicia Meusel[1] · Nadine Scheller[1] · Günter Daniel Rey[1] · Sascha Schneider[2]

Received: 28 September 2023 / Accepted: 9 January 2024 / Published online: 7 February 2024

Abstract
Color has been investigated as a signaling cue in multimedia learning environments, guiding the learner's attention and as an emotional design element, increasing the learner's motivation and, thus, improving learning outcomes. Retrieval cues (e.g., visual cues, odor, sound) facilitating memory retrieval have been primarily investigated in learning simple word lists. Contrary to additional retrieval cues, the background color is a component that is always present in multimedia learning environments. This study investigates if the background color of learning texts as retrieval cue can enhance learning and affect cognitive load and motivation. Hypotheses are formulated according to the *Cognitive-Affective Theory of Learning with Multimedia* (CATLM) and in the setting of context-dependent memory, specifically the *Information, associated Context, and Ensemble Theory* (ICE). A 2 (related vs. unrelated background color) × 2 (with vs. without colors in the learning test) -factorial between-subjects design with an additional control group was utilized. For the control group, the background of the learning texts and learning questions was white. In total, 191 native German speakers were randomly assigned to the five groups. The findings indicate that relying solely on the background color as a retrieval cue is insufficient. Instead of facilitating memory retrieval, the background color remains context information stored separately from the item information. However, the results should be approached carefully as learning outcomes may be subject to ceiling and floor effects.

Keywords Retrieval cues · Context-dependent memory · Multimedia learning · Cognitive load · Background color

Felicia Meusel and Nadine Scheller contributed equally to this work.

Extended author information available on the last page of the article

1 Introduction

Today, *Information and Communication Technology* (ICT) is omnipresent. Information and communication technologies are, for example, new multimedia technologies, including: “computer software, CD-ROM, the internet, television, film as well as internet-based project work, e-mail, chat, blogs, wikis, podcasts, and so forth” (Zhou, 2018, p. 22). Education is another area of our lives that benefits from multimedia tools. Therefore, several learning theories have been formulated in recent decades that focus on different aspects of (multimedia) learning. One of these theories is the *Cognitive-Affective-Social Theory of Learning in Digital Environments* (CASTLE) by Schneider et al. (2022). The CASTLE suggests that there are four kinds of processes that influence learning outcomes. Cognitive processes rely on cognitive factors like attention, working memory capacity, and cognitive load. Affective processes are influenced by changes in motivation, interest, and emotion. Social processes rely on social cues, leading learners to perceive learning situations as social interactions and thus activating schemata and rules of social interactions. Meta-cognitive processes are based on knowledge about mental states and cognition. These kinds of processes are intertwined and mutually influence each other. Based on these processes, research in multimedia learning investigates how instructors can design their learning materials to facilitate and improve learning (Mayer & Fiorella, 2021). For instructors, knowing which design elements affect learning outcomes and how and on which levels of processing these design elements affect learning is essential. This knowledge allows instructors to design and create suitable information and communication technologies for their courses and classes. Color is a prominent design element affecting cognitive and affective processes in learning. Many studies investigated the effects of colors in multimedia learning environments on learning outcomes and learner motivation (for an overview, see Alpizar et al., 2020; Wong & Adesope, 2021). Color is thought to increase learner motivation by making the learning environment visually more appealing and to direct the learner’s visual attention during the learning process. However, it is still unclear if color can affect learning in other ways than increasing motivation and guiding attention and whether the specific choice of colors matters.

1.1 Theoretical background

1.1.1 Color in multimedia learning

Color as an affective factor According to the *Cognitive-Affective Theory of Learning with Multimedia* (CATLM; Moreno & Mayer, 2007), learning is influenced by cognitive, meta-cognitive, and affective factors. Affective factors influence learning through changes in motivation and emotions, which can increase or decrease the learner’s cognitive engagement. Making a learning environment visually more appealing is supposed to increase the learner’s motivation to engage in profound and meaningful learning and thus improve learning outcomes. A meta-analysis by Wong and Adesope (2021) on *emotional design* features, including color, shapes, and *anthropomorphism* (i.e., the attribution of human-like features to non-human objects; Wong & Adesope, 2021), found significant beneficial effects of emotional design in multimedia learning environments on *retention* (the memory of key facts; Brom et al., 2018), *transfer* (the ability to apply the learned concepts in new situations and to new problems; Brom et al., 2018), *comprehension* (the understanding of key concepts; Brom et al., 2018), positive affect, intrinsic motivation, mental effort, liking/enjoyment, and perceived difficulty. For color alone, only effects on retention and intrinsic motivation were found (Wong & Adesope, 2021). A previous meta-analysis by Brom et al. (2018) on emotional design found significant positive effects of anthropomorphism and/or pleasant colors on retention, comprehension, transfer, and intrinsic motivation. Effects on liking/enjoyment were weak, while only marginal effects on positive affect were found. However, Brom et al. (2018) did not differentiate between the effects of anthropomorphism, color, or a combination of both on learning and motivation. Other studies also found beneficial effects of utilizing positive emotional designs as a combination of color and anthropomorphism on retention and transfer (e.g., Li et al., 2020; Peng et al., 2021), but no significant difference in emotion (e.g., Li et al., 2020; Shangguan et al., 2020) nor in motivation (e.g., Shangguan et al., 2020). Shangguan et al. (2020) further found that using colorful design with anthropomorphism increased learners' perceived difficulty and mental effort but had no effect on retention and transfer. However, these studies investigated the effects of a positive emotional design as a combination of color and anthropomorphism. Therefore, whether the effects arise from color, anthropomorphism, or a combination of the two remains unclear. Studies separately investigating the effects of color on learning found that a colorful chromatic design positively impacted recall and mental effort when compared to an achromatic design (Uzun & Yıldırım, 2018). Additionally, transfer performance (Plass et al., 2014; Wang et al., 2023) and comprehension (Plass et al., 2014) were positively affected by a chromatic design. However, no significant effect on motivation was found when using color alone (Plass et al., 2014).

Color as a cognitive factor Color can also affect learning at the level of cognitive processing. Research on the cognitive effects of color in multimedia learning has focused on using color as a signaling cue by highlighting the essential elements of the learning material and thus guiding the learner's attention to them (e.g., Alpizar et al., 2020; Mayer, 2002). According to the *Cognitive Load Theory* (CLT; Sweller et al., 1998, 2019), three kinds of cognitive load are involved in information processing during learning. *Intrinsic cognitive load* (ICL) refers to task-relevant load caused by the complexity of the learning topic. This complexity is determined by the element interactivity and the learner's prior knowledge. *Extraneous cognitive load* (ECL) is the task-irrelevant cognitive load that increases by adding task-irrelevant information to the learning environment. *Germane cognitive load* (GCL) depends on constructing and automating mental schemata into long-term memory. By guiding the learner's attention to the essential elements of the learning material, signaling can facilitate the selection of relevant information and, as a result, reduce the cognitive load (Alpizar et al., 2020; Mayer, 2002).

Skulmowski (2022) investigated the effects of color coding of the to-be-learned information on learning outcomes. Participants were taught the parts of a fictional bone with either a colored or an achromatic graphic of the bone. Color coding was beneficial for learning when used both during learning and during retrieval. However, color coding was detrimental to learning when used only in one of the two phases. This study focused on the segmenting function of the colors. The fictional bone consisted of several components, each with a different color. Thus, the colors segmented the visualization of the bone into visual chunks, making them easier to distinguish and thus reducing the cognitive load (Skulmowski, 2022).

However, the colors utilized in Skulmowski's study (2022) can also be seen as memory cues, as each color could be assigned to one part of the fictional bone. Other studies investigated the role of color as a memory cue when used as environmental information: Clariana and Prestera (2009) examined the effects of a colored left margin on memory retrieval, and Prestera et al. (2005) investigated the effects of colored borders on memory retrieval. The studies found that color had only a small or no effect on memory retrieval when color was used as environmental information.

1.1.2 Context-dependent memory

Using color as environmental information to facilitate learning is based on the theory of context-dependent memory (Smith & Vela, 2001). The context-dependent memory theory states that memory retrieval is facilitated when the context in which the information was learned is reinstated at the time of memory retrieval. Godden and Baddeley (1975) found that scuba divers who learned two separate word lists, one on land and the other underwater, remembered the words from the list learned on land better when retrieved on land. The words from the list learned underwater were remembered better when retrieval occurred underwater, too. Overall, in a meta-analysis on context-dependent memory by Smith and Vela (2001), the effect of context on memory retrieval was found to be small but significant ($d = 0.28$). However, the authors emphasized that many studies found large context effects. The mechanisms that underlie context-dependent memory can be explained by the *Information, associated Context, and Ensemble Theory* (ICE; Murnane et al., 1999). The ICE proposes that three kinds of information are involved in context-dependent recognition: item information, context information, and ensemble information. Item information is task-relevant and, therefore, is primarily encoded. Context information includes any environmental information that is task-irrelevant. When item and context information are processed, they merge into ensemble information, a combination of the item and context information. Whenever the context information is combined with the item information into ensemble information, the context information can serve as a retrieval cue to increase memory performance. Presenting the context information again activates the stored ensemble information in memory, thereby accessing the item information associated with the ensemble. However, if the context information fails to become part of the combined ensemble information, it cannot be used as a retrieval cue. Whether context information becomes part of the combined ensemble may rely on learner characteristics, the instruction explicitly prompting the learner to build ensembles or the relationship between context and item information regarding the content.

1.1.3 Background color as a retrieval cue

The background color is context information always present in the learner's visual field (Isarida & Isarida, 2007), regardless of whether the learner is aware of the background. Previous studies investigated the effects of background color in print and digital media on readability and legibility (e.g., Zorko et al., 2017) as well as on memory and learning performance (e.g., Acharya et al., 2021; Hsieh et al., 2022; Jadhao et al., 2020; Sattarzadeh & Tahmasebi Boroujeni, 2021; Sun & Liu, 2016). A review of the effects of color on memory performance by Dzulkifli and Mustafar (2013) states that color can impact memory performance due to changes in attention and arousal levels. However, some studies indicate that colors can increase memory performance, while others indicate the opposite (Dzulkifli & Mustafar, 2013). Regardless of the effects of different background colors on memory, the background color may enhance memory retrieval as a retrieval cue if it becomes part of the ensemble information. Context information can be automatically integrated into the ensemble information (due to learner characteristics) or if the instruction contains the note that the background color may help remember the information better (Murnane et al., 1999). However, studies on the effectiveness of retrieval cues primarily investigated odor (e.g., Parker et al., 2001), music (e.g., Hwang & Kim, 2015), the speaker's voice (e.g., Dodson & Shimamura, 2000), realism (e.g., Skulmowski & Rey, 2021), and the general visual background (e.g., Chung & Cheon, 2020; Smith & Manzano, 2010). Only a few studies investigated background color as a retrieval cue (Isarida & Isarida, 2007; Pointer & Bond, 1998; Rutherford, 2004; Sakai et al., 2010), resulting in mixed findings. Also, most studies have focused on the effect of retrieval cues on learning simple word lists (e.g., Godden & Baddeley, 1975; Isarida & Isarida, 2007; Parker et al., 2001; Smith & Manzano, 2010; Tulving & Pearlstone, 1966). Retrieval cues rarely have been investigated in multimedia learning environments, including entire learning texts (e.g., in Clariana & Prestera, 2009; Prestera et al., 2005). However, in these studies, only parts of the background were colored. According to Karpicke (2012), who investigated retrieval-based learning, the critical factor for all learning is the diagnostic value of retrieval cues. The diagnostic value is the degree to which a specific cue helps the learner recover particular target knowledge while excluding competing candidates (Karpicke, 2012).

The findings on using color as a retrieval cue in multimedia learning environments are mixed. While Clariana and Prestera (2009) and Prestera et al. (2005) found no or small effects of color in the background, Skulmowski (2022) found a beneficial effect of color coding when used both during learning and recall processes.

The present study aims to investigate using background color as a retrieval cue in multimedia learning. According to the ICE (Murnane et al., 1999), learners can integrate item and context information to form ensemble information when prompted that the context information may help them remember the items. Thus, a retrieval cue effect of the background color of learning texts on retention and transfer performance was expected when the instructions included the prompt that each text has a different background color that may help to remember the content of the text. Additionally, choosing a background color related to the content of the learning text may further increase the likelihood of the background color becoming part of the combined ensemble information. The retrieval cue effect of background color on learning outcomes may increase for colors related to the content of the learning texts compared to background colors unrelated to the content. Information on how blue pigment was gained from natural sources should be remembered better when presented and retrieved on a brown background than when presented on a white background. The effect should be even more prominent when the text about gaining natural blue pigment is presented and retrieved on a blue background instead of a brown background. Retrieval cue effects should only occur when the same background color is present during learning and testing. According to the ICE, the following hypotheses were formulated:

H1: Learners score higher on retention and transfer tests when the background color of the learning questions matches the color of the background of the learning texts than when the background of the learning texts is colored and the background of the learning questions is white.
H2: Learners score higher on retention and transfer tests when the background color of the learning texts and learning questions matches the content of the learning text than when the background color of the learning texts and learning questions does not match the content of the learning text.

In line with the CATLM, the following hypotheses were formulated:

H3: Learners are more motivated when the background of the learning texts and learning questions is colored than when the background of the learning texts and learning questions is white.
H4: Learners score higher in GCL ratings when the background of the learning texts and learning questions is colored than when the background of the learning texts and learning questions is white.

2 Method

2.1 Design

The experiment utilized a 2×2-factorial between-subjects design (see Fig. 1) with the factors "background color of the learning texts" (related to text vs. unrelated to text) and "background color of the learning questions" (colored vs. white). The learning texts were presented on a colored background in all four experimental groups. The background color of the learning texts either corresponded to the content of the text (e.g., blue background + text about the color blue) in the related-to-text condition or did not correspond to the content of the text (e.g., brown

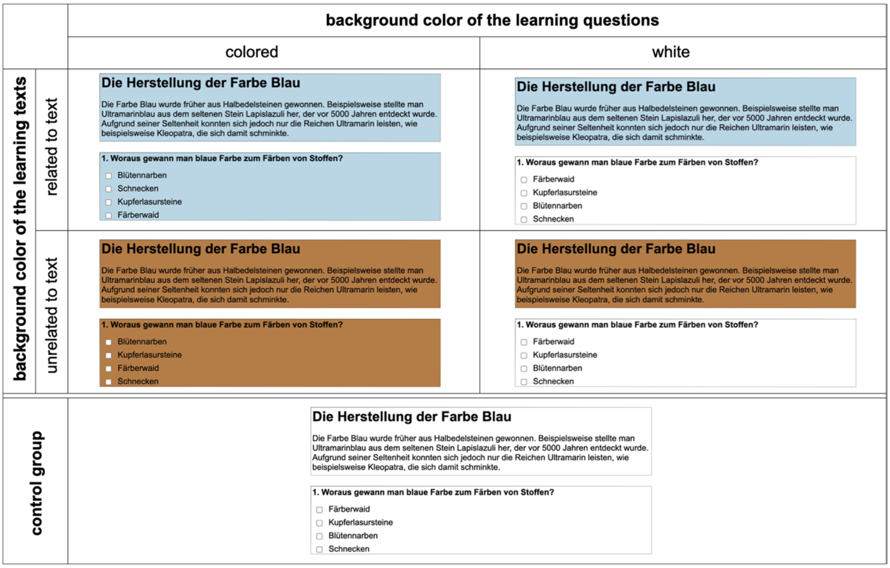


Fig. 1 Experimental design. *Note.* In the learning environment, the backgrounds of the learning texts and the learning questions either were colored and related to the text (ReC; top left) or were colored and unrelated to the text (UnC; bottom left) or the background of the learning texts was related to the text while the background of the learning questions was white (ReW; top right) or the background of the learning texts was unrelated to the text while the background of the learning questions was white (UnW; bottom right). In the control group, all survey pages were white with black font

background + text about the color blue) in the unrelated-to-text condition. The background colors for the unrelated-to-text condition were randomly assigned to the four texts. They were the same for all participants in the two experimental groups with unrelated background colors. The background of the learning questions was either white or colored. In the groups where the backgrounds of the learning questions were colored, the backgrounds were the same color as those of the corresponding learning text. In the groups where the learning questions had a white background, the learning texts had a colored background. The text order was randomized for each participant to avoid order effects. In an additional control group, the background of the learning text, the background of the learning questions, and the background throughout the rest of the survey were white. In the four experimental groups, the background of the survey pages not showing the learning text or the learning questions was white.

2.2 Participants

An a priori power analysis conducted with G*Power 3.1.9.7 (Faul et al., 2007) yielded a minimum sample size of 160 participants for $f = 0.25$, $\alpha = 0.05$, and $1 - \beta = 0.80$. After the ethics committee of the Chemnitz University of Technology approved the study, 200 native German speakers – recruited on Prolific (www.

prolific.co) [March 2023] – participated in the study. Nine participants had to be excluded from the data analysis. Seven participants showed signs of color vision deficiency, and two did not participate conscientiously in this study. The remaining 191 participants were 18 to 63 years old ($M = 29.51$, $SD = 8.35$). Eighty-nine participants indicated to be male, 99 were female, one participant identified as diverse, and two participants did not want to specify. Of the 191 participants, 90 were employed, 73 were students, 13 were job-seekers, four were trainees, three were unemployed, two were secondary school students, and six categorized as others (see Table 2 in Appendix 1). All participants were randomly assigned to one of five conditions and received £4.50 for completing the study.

2.3 Materials

In an online survey, participants read learning texts about how the colors blue, green, yellow, and brown were obtained naturally in the past. The online survey was created using SoSci Survey (Leiner, 2023) and was accessible to participants through www.soscisurvey.de. The learning text was divided into four parts. Each part addressed the production methods of one of the four color pigments (blue/green/yellow/brown). For example, the text on how blue pigment was obtained naturally in the past included the information that blue color pigment used to be extracted from semiprecious stones (lapis lazuli) and only the rich could afford ultramarine, such as Cleopatra, who used it for make-up. Additionally, the text specified how blue pigment could be gained from copper glaze stones and plants like woad, how the pigment was extracted from these resources, and what the differently obtained color pigments were helpful for. Similar information was provided for the colors green, yellow, and brown. The four texts comprised 209 to 249 words ($M = 225.75$, $SD = 16.76$). Each text was presented on a separate page in the center and formatted in a justified font. The font was the same in all five groups. The headings were displayed in bold and separated from the text by a blank line. The font color was always black, while the background color varied depending on the experimental condition (see Fig. 2 and Table 3 in Appendix 2). The entire page was colored in the corresponding background color (see Fig. 3 in Appendix 2). In the control group, the font color was black, and the background of every survey page was white (see Fig. 1). The text order was randomized. Participants could complete the survey at their own pace by clicking a "continue" button (see Fig. 3 in Appendix 2).

2.4 Procedure

Participants received written instructions. By starting the study in Prolific, participants were redirected to the study in SoSci Survey. Participants were randomly assigned to one of the four experimental conditions or to the control group. The groups were distributed equally by using a random generator from SoSci Survey. After reading the privacy policy, the participants had to agree consciously to participate in the study. The survey started by collecting demographic information from the participants (e.g., age, gender, first language). Next, color vision deficiency was checked using eight color plates from the Ishihara color test (Ishihara, 1917). Before the learning phase, the participants' prior knowledge of the subject was measured with three open-ended questions. Participants in the experimental groups were told that each text had a different background color, which could help them remember the information. In the control group, this prompt was excluded from the instruction to avoid confusion. After reading the note about the potential usefulness of the background colors, the participants read all four texts about the different colors at their own pace (reading time in minutes: $M = 5.53$, $SD = 3.20$). As mentioned above, the four texts were presented in randomized order to prevent order effects. After reading the four texts, participants responded to eight multiple-choice retention questions (two questions per color) and eight open-ended transfer questions (two questions per color). The questions were presented separately for each color to ensure that the background matched the background of the corresponding learning text. Subsequently, the participants' cognitive load and motivation were measured. Ultimately, the participants were thanked for their participation and redirected to Prolific to confirm participation and receive compensation for their efforts. All participants completed the study in approximately 30 min.

2.5 Measures

Color vision deficiency Color vision deficiency was checked using eight color plates from the Ishihara color test (Ishihara, 1917). Plates one, two, six, eight, 12, 14, 15, and 16 were chosen. This selection covered all types of color vision deficiency that the Ishihara color test can measure. The plates show a circle with a number in one color on a background of a different color. The number and the background consist of multiple dots of different sizes and slightly different color tones. Figure 4 in Appendix 2 shows one of the plates.

Prior knowledge Prior knowledge was assessed through three open-ended questions (e.g., Which methods for extracting color pigments from natural resources do you know?), rated by two independent raters. Participants could score between zero and two points per question, resulting in a maximum score of six points. Inter-rater reliability between the two raters was calculated using Cohen's Kappa, resulting in a score of $\kappa = 0.67$, $p < 0.001$. Prior knowledge among all participants was low ($M = 2.31$, $SD = 1.26$) and did not differ between the five groups, $F(4, 186) = 0.12$, $p = 0.98$. The internal consistency value (Cronbach's α) for prior knowledge was low, $\alpha = 0.38$.

Learning outcomes Using retention and transfer performance tests, learning outcomes were assessed as a dependent variable. Retention performance was measured using eight multiple-choice questions (two questions per color; e.g., From which natural resources can you get blue pigment?). For each question, four possible answers were given in random order. Between one and four answers per question were correct. Participants could score up to four points per question (one point for each correctly marked target answer and correctly unmarked distractor answer), resulting in a maximum of 32 points. Transfer performance was measured using eight open-ended questions (two questions per color; e.g., A painting from the Middle Ages is found, but the trees and bushes are faded. Why?). Two raters independently compared the answers to the transfer questions to predefined correct answers. The participants could reach a maximum of two points per transfer question, resulting in a maximum of 16 points. The inter-rater reliability score between the two raters was a Cohen's Kappa of $\kappa = 0.70$, $p < 0.001$. Internal consistency values for the learning outcomes are low, with $\alpha = 0.44$ for retention and $\alpha = 0.22$ for transfer.

For retention and transfer measurement, both questions for each color were presented together on a single page, resulting in four pages for the retention and four for the transfer questions. The individual pages of the retention and transfer tests were presented in random order for each participant. First, participants had to answer all eight retention questions and, afterward, all eight transfer questions.

Cognitive load The cognitive load questionnaire from Klepsch et al. (2017) measured cognitive load as a dependent variable. The questionnaire measures the three sub-dimensions of cognitive load (intrinsic, extraneous, and germane cognitive load) with seven items that are rated on a seven-point Likert scale ranging from 1 = "completely disagree" to 7 = "completely agree." Two items measure the intrinsic cognitive load (ICL; e.g., For this task, many things needed to be kept in mind simultaneously.), two items measure the germane cognitive load (GCL; e.g., I made an effort, not only to understand several details, but to understand the overall context.), and three items measure the extraneous cognitive load (ECL; e.g., During this task, it was difficult to recognize and link the crucial information.). Regarding internal consistency, Klepsch et al. (2017) report mean Cronbach's α values for the ICL scale ($\alpha = 0.81$), ECL scale ($\alpha = 0.86$), and the GCL scale ($\alpha = 0.67$).

Motivation The participants' motivation was assessed as a dependent variable using a shortened version of the *situational motivation scale* (SIMS; Guay et al., 2000). The participants answered the question "Why did you do the previous activity?" on a seven-point Likert scale ranging from 1 = "completely disagree" to 7 = "completely agree." Therefore, participants rated four items on the sub-scale of intrinsic motivation (e.g., Because I think that this activity is interesting.) and four items on the sub-scale of external regulation (e.g., Because I am supposed to do it). The sub-scale of intrinsic motivation demonstrated internal consistency (Cronbach's α) of $\alpha = 0.95$, and the sub-scale of external regulation demonstrated internal consistency (Cronbach's α) of $\alpha = 0.86$.

3 Results

Statistics Statistical analysis was performed in IBM SPSS Statistics (Version 28). A 2×2-factorial multivariate analysis of covariance (MANCOVA) was performed with prior knowledge as a covariate. Retention and transfer performance, motivation, and cognitive load in terms of GCL were included as the dependent variables. The relation of the background color of the learning texts to the contents of the learning text and the background of the learning questions were included as the independent variables. Prior knowledge was included as a covariate because changes in the learner's prior knowledge change

ICL, and ICL affects learning outcomes (Sweller et al., 2019). Also, prior knowledge has been shown to affect the effects of design principles on learning outcomes (expertise reversal effect; Kalyuga, 2007). A significant effect of prior knowledge was found on retention performance, $F(3, 154)=7.56$, $p=0.01$, $\eta_p^2=0.05$, ICL, $F(3, 154)=5.28$, $p=0.02$, $\eta_p^2=0.03$, and ECL, $F(3, 154)=3.97$, $p=0.05$, $\eta_p^2=0.03$.

A one-factorial MANCOVA was performed with prior knowledge as the covariate to compare the experimental groups with the control group. The dependent variables included retention performance, transfer performance, motivation, and GCL scores. The group was included as the independent variable. Again, prior knowledge was included as a covariate. A significant effect of prior knowledge on retention performance, $F(4, 186)=5.65$, $p=0.02$, $\eta_p^2=0.03$, and transfer performance, $F(4, 186)=5.58$, $p=0.02$, $\eta_p^2=0.03$, was found.

Learning outcomes Retention performance did not differ between the five groups, $F(4, 186)=0.06$, $p=0.99$ (see Table 1 and Fig. 5 in Appendix 3). The main effect of the relation of the background color of the learning text to the content of the learning text, $F(3, 154)=0.09$, $p=0.77$, the main effect of the background of the learning questions, $F(3, 154)=0.01$, $p=0.91$, and the interaction of the two factors, $F(3, 154)=0.03$, $p=0.87$, did not become significant.

Transfer performance did not differ between the five groups, $F(4, 186)=1.10$, $p=0.36$ (see Table 1 and Fig. 6 in Appendix 3). The main effect of the relation of the background color of the learning text to the content of the learning text, $F(3, 154)=0.20$, $p=0.66$, the main effect of the background of the learning questions, $F(3, 154)=0.61$, $p=0.44$, and the interaction of the two factors, $F(3, 154)=1.11$, $p=0.29$, did not become significant.

Cognitive load GCL scores did not differ between the five groups, $F(4, 186)=0.32$, $p=0.87$. The main effect of the relation of the background color of the learning text to the content of the learning text, $F(3, 154)=0.54$, $p=0.47$, the main effect of the background of the learning questions, $F(3, 154)=0.63$, $p=0.43$, and the interaction of the two factors, $F(3, 154)=0.08$, $p=0.78$, did not become significant.

Table 1 Means and standard deviations of the retention and transfer scores per group

Learning measure	ReC		UnC		ReW		UnW		Control	
	Mean	*SD*	Mean	*SD*	Mean	*SD*	Mean	*SD*	Mean	*SD*
Retention	25.25	3.46	25.32	3.73	25.17	3.22	25.41	3.32	25.12	3.38
Transfer	4.75	2.69	4.53	2.22	4.11	1.96	4.65	2.27	3.86	2.34

This table demonstrates the mean scores and standard deviations of the two learning measures for the four experimental groups: "the background of the learning text is related to the text, the background of the learning questions is colored" (ReC), "the background of the learning text is unrelated to the text, the background of the learning questions is colored" (UnC), "the background of the learning text is related to the text, the background of the learning questions is white" (ReW), "the background of the learning text is unrelated to the text, the background of the learning questions is white" (UnW), and the control group (control)

Motivation Motivation scores were not significantly different between the five groups, $F(4, 186) = 0.56$, $p = 0.69$. The main effect of the relation of the background color of the learning text to the content of the learning text, $F(3, 154) = 0.39$, $p = 0.53$, the main effect of the background of the learning questions, $F(3, 154) = 0.01$, $p = 0.91$, and the interaction of the two factors, $F(3, 154) = 0.06$, $p = 0.82$, did not become significant.

4 Discussion

Learning outcomes The present study did not find a retrieval cue effect of background color on learning outcomes when learning with texts instead of simple word lists. However, this may be due to an ill-chosen measure of learning outcomes. Differences in learning outcomes between the five groups may not have reached the significance level due to a ceiling effect. Participants could achieve one to 32 points in the retention test, and all groups had a mean of around 25 points (min = 18.00, max = 32.00; see Table 1). For the transfer performance, however, the results indicate a floor effect. Participants could achieve one to 16 points, and all groups reached around four points (min = 0.00, max = 12.00; see Table 1). Only one person scored 12 points on the transfer test and must be treated as an outlier, as seen in Fig. 6 in Appendix 3. This suggests that the retention test might have been too easy, while the transfer test may have been too challenging. During the transfer test, some participants expressed their inability to answer the questions as the learning text did not directly provide the answers. In the study, the participants were not instructed that they would have to transfer the knowledge they gained from the text to new problems to answer the transfer questions. Thus, future studies should inform the participants that some questions cannot be directly answered with information from the text but require the participants to transfer their knowledge to new problems. Additionally, the relation of the background color to the content of the learning text did not influence retention and transfer scores. Using background colors that relate to the content of the learning text does not make the background color become part of the ensemble information. Thus, hypotheses H1 and H2 – formulated in line with the ICE – could not be confirmed. This indicates that when designing information and communication technologies such as learning web pages, instructors do not have to think about the background color of the learning environments. Different background colors can have different effects on memory (Dzulkifli & Mustafar, 2013) but information and communication technology designers should not put too much effort in using background colors as retrieval cues. However, results may differ for an appropriate learning outcome measure without ceiling and floor effects.

Another possibility may be that the background color alone is insufficient as a retrieval cue. At the beginning of the study, the participants received the information that each text had a different background color that could help them remember the information. However, it is possible that the participants did not pay enough attention to the background color or did not integrate it with the text's contents. The background color as context information may not have become part of the ensemble information. This could explain why the relation of the background color to the texts' contents did not affect learning outcomes. Merkt et al. (2020) investigated if the background in a learning video could affect learning and serve as a retrieval cue. The video's background depicted either a topic-related authentic scenery (such as an image of plants in a greenhouse while teaching about plants) or a neutral gray background. During the learning test, the participants were presented with either the same background or the other one. The study found no retrieval cue effect on learning outcomes, and the relationship between the background and the topic did not affect learning outcomes. However, participants were not informed about the background, which was the same throughout the video. Without an explicit prompt, even a complex matching background as context information did not become part of the ensemble information. Further studies are required to investigate if the background in a learning environment can serve as a retrieval cue in learning with entire texts and how the background can be salient enough to become part of the ensemble information and serve as a retrieval cue.

Additionally, in contrast to prior studies (e.g., Clariana & Prestera, 2009), the participants' age range was very heterogeneous, ranging from 18 to 63 years. This may also be a reason why there were less significant results. Furthermore, a more scholarly population was likely obtained by conducting the study on Prolific, suggesting that retrieval cues may not have been needed in participants' memory recall (due to ceiling effects). Further research is needed to determine if younger participants (e.g., primary school students) or educationally disadvantaged people would benefit more from using background color as a retrieval cue.

Furthermore, the results in learning performance may have been confounded by the effects of the different background colors on memory. A review of the effects of colors on memory by Dzulkifli and Mustafar (2013) reports that colors may affect memory performance. However, the evidence is inconsistent. Cool background colors (blue) have been shown to increase short-term memory recall compared to warm background colors (red; e.g., Sattarzadeh & Tahmasebi Boroujeni, 2021). Other studies found that warm background colors increase learning performance and short-term memory compared to cool background colors (e.g., Hsieh et al., 2022; Jadhao et al., 2020; Sun & Liu, 2016). Other studies found no effect of a colored background compared to a white background (e.g., Acharya et al., 2021) regarding retention and recall. Red was excluded as a background color from the present study because it is associated with danger and errors (Jadhao et al., 2020). However, blue was included as a background color in the study. Future research should investigate if different background colors (e.g., blue, yellow, green, and brown) affect memory and learning performance differently.

Motivation Contrary to hypothesis H3 and the CATLM, the learners' motivation did not increase when adding a colored background to the learning text. This indicates that background color alone does not make a learning environment visually appealing enough to motivate learners. Other design elements are required to increase learners' motivation. The background color alone may not have been salient enough to increase motivation.

Cognitive load The results suggest that using a colored background in multimedia learning environments does not increase the GCL compared to a white background. These results do not align with the CATLM, and hypothesis H4 was not supported. This may be due to a lack of motivation difference between the experimental and control groups. Participants did not invest more resources to build mental schemata of the learning content.

Information and communication technologies The present results do not allow implications on how multimedia learning environments should be designed to facilitate and improve learning. A meta-analysis on the effects of digital technology on deep learning by Wu (2023), however, found that digital technology significantly positively affects learning outcomes in comparison to traditional learning methods. Thus, ICT bears a large potential to foster deep learning. Future studies should further investigate the effects of specific design elements of ICT (such as background color) on learning outcomes.

5 Limitations and future directions

The present study has several limitations. First, measuring the cognitive load after the learning test may have resulted in scores representing the overall cognitive load rather than the load caused by the learning intervention. Future studies should measure the cognitive load immediately after the learning intervention to avoid this. Second, motivation was also measured only after the learning test. Thus, motivation scores may reflect motivation related to the learning test rather than motivation based on the learning material. In future studies, motivation should be measured immediately after the learning intervention. Moreover, motivation should be measured before the learning intervention to determine the learners' baseline motivation and to investigate whether the learning environment increases or decreases learner motivation. Third, no conclusions regarding the effects of content-related background color on learning outcomes can be drawn from the results, as the learning test results may be subject to ceiling and floor effects. Further research is necessary to examine the effects of content-relevant background colors on learning outcomes with other learning topics and more suitable learning tests. Another limitation of the present study is that in the groups where the background color was unrelated to the content of the learning text, the background color of the four learning texts was the same for all participants. Future studies should randomize the unrelated background colors for each participant individually. Also, as mentioned before, the background color might not have been salient enough. The participants may have missed the reference to the benefit of the background color by skimming the instructions. Furthermore, color is an omnipresent perceptual stimulus often viewed from an aesthetic perspective (Elliot & Maier, 2014). Therefore, colors may have varied psychological and cultural effects. Color perception happens in the brain and not in the outside world. Furthermore, other languages use different words to describe the color blue (Wierzbicka, 1990). Thus, Sahlins (1976) argues that colors and their meaning are culturally influenced. This cultural impact is researched in various scientific disciplines, such as marketing (e.g., Broeder & Scherp, 2018; Cyr et al., 2010). In their study, Cyr et al. (2010) examined how different cultures perceive color in website design, finding that specific colors can evoke emotions and behaviors that vary between cultures. In general, the potential of color to elicit emotions (e.g., Cyr et al., 2010) applies to design effects such as emotional design and learning contexts, too. Future studies should investigate whether background color can be a reliable retrieval cue in different cultures.

6 Conclusion

The present study investigated whether background color can serve as a retrieval cue in multimedia learning environments, leading to enhanced learning outcomes for complete text-based learning. Further, it was examined whether the relation of the background color to the content of the learning text influences this effect. The results indicate that more than background color is needed to serve as a retrieval cue. The background color remains context information instead of becoming part of the ensemble information and is stored separately from the item information. Thus, the relation of the background color to the content of the learning text does not matter. However, these results should be interpreted with caution due to potential ceiling and floor effects on learning outcomes. More complex designs are necessary to increase learners' motivation and thereby increase GCL and learning outcomes than simply using a colored background instead of a white one.

Appendix 1

Demographics

Table 2 Demographic data

Participant number	Overall:		$N=191$
	ReC:		$n=40$
	ReW:		$n=41$
	UnC:		$n=38$
	UnW:		$n=39$
	Control:		$n=33$
Participant age	Range: 18 – 63 years	$M=29.51$	$SD=8.35$
Participant gender	Female:	51.83%	$n=99$
	Male:	46.60%	$n=89$
	Diverse:	0.52%	$n=1$
	Not specified:	1.05%	$n=2$
Participant status	Employed:	47.12%	$n=90$
	Student:	38.22%	$n=73$
	Job-seeking:	6.81%	$n=13$
	Trainee:	2.01%	$n=4$
	Unemployed:	1.57%	$n=3$
	Secondary school student:	1.05%	$n=2$
	Other:	3.14%	$n=6$

This table shows the number of participants in each of the five groups and the demographic data of the participants including age, gender distribution, and status

Appendix 2

Material


Die Herstellung der Farbe Blau

Die Farbe Blau wurde früher aus Halbedelsteinen gewonnen. Beispielsweise stellte man Ultramarinblau aus dem seltenen Stein Lapislazuli her, der vor 5000 Jahren entdeckt wurde. Aufgrund seiner Seltenheit konnten sich jedoch nur die Reichen Ultramarin leisten, wie beispielsweise Kleopatra, die sich damit schminkte.

Eine andere Möglichkeit, Blau zu gewinnen, waren Kupferlasursteine, die das Mineral Azurit enthalten. Azurit wird mit Granit zermahlen und anschließend fein gesiebt. Danach wird das Azurit beim "Schlämmen" mit Wasser gespült, um ein möglichst reines Blau zu erhalten. Das reine Azurit wird dann getrocknet und kann mit einem Bindemittel zu Aquarellfarbe verarbeitet werden. Als Bindemittel verwendeten die Menschen Eiweiße, Harze, Öl oder eine Mischung aus Eiweiß, Öl und Wasser.

Die Herstellung der Farbe Grün

Seit Jahrhunderten versuchten Menschen, ein echt grünes Pigment herzustellen. Im alten Ägypten war Grün das Symbol für Regeneration und Wiedergeburt. Die Farbe Grün wurde ursprünglich aus Malachit – einem Kupfermineral – gewonnen. Man versuchte das Malachit für die Bemalung von Grabwänden zu verwenden, doch es war teuer und wurde mit der Zeit dunkler.

Die alten Römer kamen auf die Idee, Kupferplatten in Wein zu tränken, um Grünspan zu erzeugen. Denselben grünen Farbton findet man heute auf patinierten Metalldächern, alten Münzen oder Skulpturen. Die Römer verwendeten das Pigment für Mosaike, Fresken und Glasmalereien. Mittelalterliche Mönche nutzten ihn zur Färbung von Manuskripten.

Im Mittelalter gewann man grünen Farbstoff primär aus Pflanzen wie Lauch, Petersilie,

Die Herstellung der Farbe Gelb

Die Farbe Gelb wurde im Mittelalter aus der Reseda-Pflanze gewonnen. Der aus den Pflanzen gewonnene Farbstoff war ein wichtiger Gelbton zum Färben von Stoffen. Sobald Stängel, Blätter und Kapseln der Pflanze eine gelbliche Färbung annehmen, werden die Pflanzen geerntet und zu Garben zusammengebunden. Die Garben werden anschließend getrocknet und gehäckselt. Die getrockneten Garben werden dann mit Wasser aufgekocht und der Farbstoff durch ein Leinentuch aus dem Sud gefiltert.

Safran, einen besonders leuchtend gelben Farbstoff, kann man aus den Blütennarben einer violett blühenden Krokusart gewinnen. Den Farbstoff erhält man durch das Auswaschen der getrockneten Blütennarben. Für 1 kg Safran musste man 120.000 bis 200.000 Blüten sammeln. Tiergalle von Ochsen, Kälbern oder Schildkröten ergab im natürlichen Zustand oder vermischt mit

Die Herstellung der Farbe Braun

Die Farbe Braun wurde aus braunen Erden gewonnen, auch bekannt als Umbra oder Braunocker. Diese enthalten eisen- und manganhaltige Tone, welche die braune Farbe verursachen. Je höher der Mangangehalt, desto intensiver die Braunfärbung. Braune Erden findet man in Italien, Zypern, England und Kleinasien. Aber auch in Deutschland kann man braune Erden am Rhein und im Harz finden. Der Braunocker wird fein gemahlen und gesiebt, mit Wasser gereinigt und getrocknet. Das fertige Pigment kann man dann mit Bindemitteln zu einer Farbe verreiben. Auch Höhlenmalereien wurden bereits mit Farbe aus braunen Erden angefertigt.

Im Mittelalter wurde braune Farbe durch das Vermischen von gelbem Ocker mit schwarzem Ruß oder mit roter Erde hergestellt. Mittelalterliche Mönche trugen braune Roben als Zeichen ihrer Demut und Armut. Umbra ist etwa seit dem 16. Jahrhundert in der Malerei anzutreffen. Da

Fig. 2 Example of the four colored learning texts in the experimental groups with a background of the learning text that is related to the text. *Note.* This figure shows screenshots of the four learning texts. One text was about the extraction of blue pigment from natural sources (top left), one about green pigment (top right), one about yellow pigment (bottom left), and one about brown pigment (bottom right). Each text was presented on a separate page and had a different background color that was either related to the text (as can be seen here) or unrelated to the text. In the control group, the background of all four texts was white

Table 3 Hex-codes used for the background colors

color	hex-code
blue	#BAD6E5
green	#89AC76
yellow	#FFFF99
brown	#B47D49

This table shows the four colors addressed in the learning texts and used as background colors and the hex-codes used

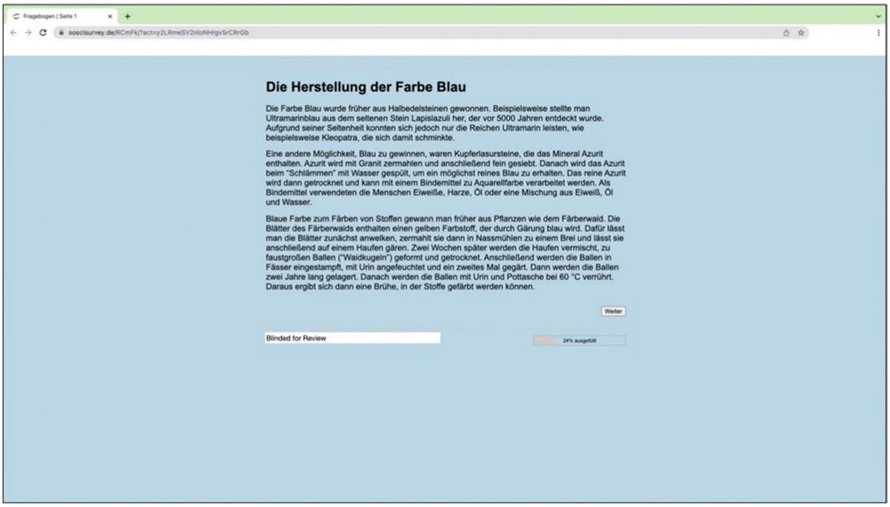
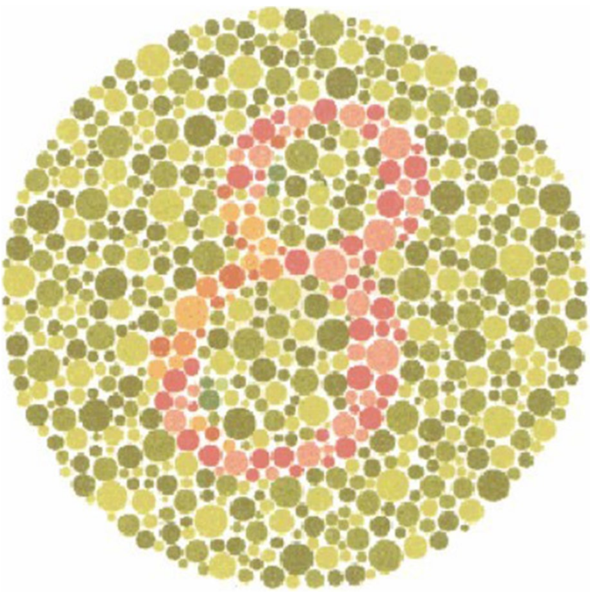


Fig. 3 Example of the text arrangement and coloring of the page. *Note.* This figure shows the entire online survey page of the learning text about blue pigment. The text was presented in the middle of the page. The entire background was colored. Participants could read the text in their own pace and continue to the next page of the survey by clicking a *continue* button

Fig. 4 Ishihara plate 2. *Note.* People with normal color vision see the number 8 while people with a congenital red-green color blindness see the number 3. From https://alexoptik.de/images/ishiharafarbtafeln.pdf



Appendix 3

Results

Fig. 5 Learning outcomes in terms of retention performance. *Note*. This figure shows the scores of the five groups in the retention test. White dots indicate the group means and black dots indicate outliers

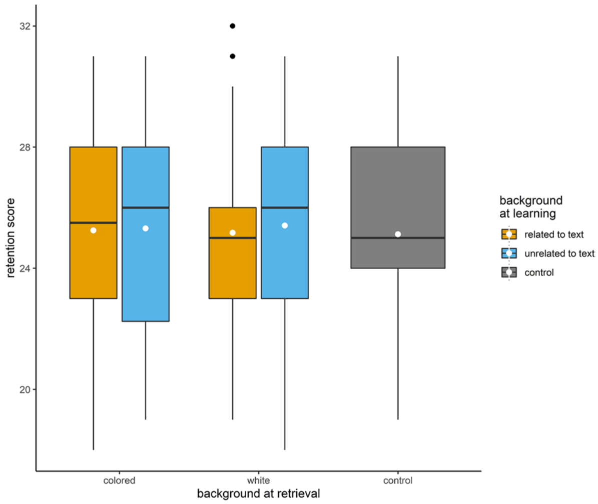
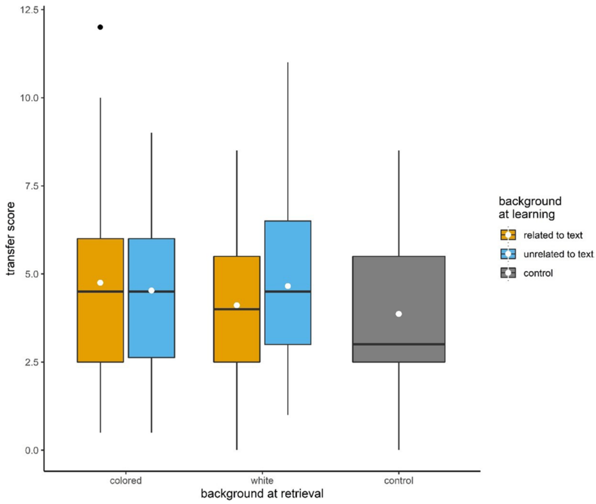


Fig. 6 Learning outcomes in terms of transfer performance. *Note*. This figure shows the scores of the five groups in the transfer test. White dots indicate the group means and black dots indicate outliers



Authors' contributions **FM** was responsible for the conceptualization, data curation, formal analysis, investigation, methodology, project administration, and writing (original draft). **NS** was responsible for the conceptualization, data curation, software, formal analysis, investigation, methodology, project administration, and writing (original draft). **GDR** was responsible for the conceptualization, supervision of the project, and writing (review and editing). **SS** was responsible for the conceptualization, supervision of the project, and writing (review and editing).

Funding This research received no specific grant from funding agencies in the public, commercial, or not-for-profit sectors. Open Access funding enabled and organized by Projekt DEAL.

Data availability The data sets of the present study are available from the corresponding author upon reasonable request.

Declarations

Ethics approval This study was approved by the ethics committee of the Chemnitz University of Technology.

Competing interests The authors declare that they have no competing interests.

References

Acharya, M. S., Keertish, N., Gowda, H. T., & Krithishree, S. (2021). Effect of white and yellow background colour on short term memory: A cross-over study. *Journal of Clinical and Diagnostic Research, 15*(6), VC07–VC09. https://doi.org/10.7860/JCDR/2021/46349.15042

Alpizar, D., Adesope, O. O., & Wong, R. M. (2020). A meta-analysis of signaling principle in multimedia learning environments. *Educational Technology Research and Development, 68*(5), 2095–2119. https://doi.org/10.1007/s11423-020-09748-7

Broeder, P., & Scherp, E. (2018). Colour preference of online consumers: A cross-cultural perspective. *Marketing from Information to Decision Journal, 1*(1), 1–11. https://doi.org/10.2478/midj-2018-0001

Brom, C., Stárková, T., & D'Mello, S. K. (2018). How effective is emotional design? A meta-analysis on facial anthropomorphisms and pleasant colors during multimedia learning. *Educational Research Review, 25*, 100–119. https://doi.org/10.1016/j.edurev.2018.09.004

Clariana, R. B., & Prestera, G. E. (2009). The effects of lesson screen background color on declarative and structural knowledge. *Journal of Educational Computing Research, 40*(3), 281–293. https://doi.org/10.2190/EC.40.3.b

Chung, S., & Cheon, J. (2020). Emotional design of multimedia learning using background images with motivational cues. *Journal of Computer Assisted Learning, 36*(6), 922–932. https://doi.org/10.1111/jcal.12450

Cyr, D., Head, M., & Larios, H. (2010). Colour appeal in website design within and across cultures: A multi-method evaluation. *International Journal of Human-Computer Studies, 68*(1–2), 1–21. https://doi.org/10.1016/j.ijhcs.2009.08.005

Dodson, C. S., & Shimamura, A. P. (2000). Differential effects of cue dependency on item and source memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 26*(4), 1023–1044. https://doi.org/10.1037/0278-7393.26.4.1023

Dzulkifli, M. A., & Mustafar, M. F. (2013). The influence of colour on memory performance: A review. *The Malaysian Journal of Medical Sciences, 20*(2), 3–9.

Elliot, A. J., & Maier, M. (2014). Color psychology: Effects of perceiving color on psychological functioning in humans. *Annual Review of Psychology, 65*(1), 95–120. https://doi.org/10.1146/annurev-psych-010213-115035

Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G*power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods, 39*(2), 175–191. https://doi.org/10.3758/BF03193146

Godden, D. R., & Baddeley, A. D. (1975). Context-dependent memory in two natural environments: On land and underwater. *British Journal of Psychology, 66*(3), 325–331. https://doi.org/10.1111/j.2044-8295.1975.tb01468.x

Guay, F., Vallerand, R. J., & Blanchard, C. (2000). On the assessment of situational intrinsic and extrinsic motivation: The situational motivation scale (SIMS). *Motivation and Emotion, 24*(3), 175–213. https://doi.org/10.1023/A:1005614228250

Hsieh, A.-Y., Lo, S.-K., & Hwang, Y. (2022). Making customers more likely to come back: The role of background colour in triggering arousal to influence memory, attitude, and patronage intention. *Electronic Commerce Research*. https://doi.org/10.1007/s10660-022-09610-3

Hwang, I., & Kim, H.-K. (2015). The effect of congruency and familiarity of background music in TV advertising on the music's role as a retrieval cue. *Asia Marketing Journal, 16*(4), 2. https://doi.org/10.15830/amj.2015.16.4.1

Isarida, T., & Isarida, T. K. (2007). Environmental context effects of background color in free recall. *Memory & Cognition, 35*(7), 1620–1629. https://doi.org/10.3758/BF03193496

Ishihara, S. (1917). *Test for color-blindness*. Handaya, Tokyo, Hongo Harukicho.

Jadhao, A., Bagade, A., Taware, G., & Bhonde, M. (2020). Effect of background color perception on attention span and short-term memory in normal students. *National Journal of Physiology, Pharmacy and Pharmacology, 10*(11), 981–984. https://doi.org/10.5455/njppp.2020.10.06162202017072020

Kalyuga, S. (2007). Expertise reversal effect and its implications for learner-tailored instruction. *Educational Psychology Review, 19*(4), 509–539. https://doi.org/10.1007/s10648-007-9054-3

Karpicke, J. D. (2012). Retrieval-based learning: Active retrieval promotes meaningful learning. *Current Directions in Psychological Science, 21*(3), 157–163. https://doi.org/10.1177/0963721412443552

Klepsch, M., Schmitz, F., & Seufert, T. (2017). Development and validation of two instruments measuring intrinsic, extraneous, and germane cognitive load. *Frontiers in Psychology, 8*, 1997. https://doi.org/10.3389/fpsyg.2017.01997

Leiner, D. J. (2023). *SoSci Survey* [Computer software]. https://www.soscisurvey.de. Accessed Mar 2023.

Li, J., Luo, C., Zhang, Q., & Shadiev, R. (2020). Can emotional design really evoke emotion in multimedia learning? *International Journal of Educational Technology in Higher Education, 17*(1), 24. https://doi.org/10.1186/s41239-020-00198-y

Mayer, R. E. (2002). Multimedia learning. In B. H. Ross (Ed.), *Psychology of Learning and Motivation* (Vol. 41, pp. 85–139). Elsevier. https://doi.org/10.1016/S0079-7421(02)80005-6

Mayer, R. E. & Fiorella, L. (Eds.). (2021). *The Cambridge Handbook of Multimedia Learning* (3rd ed.). Cambridge University Press. https://doi.org/10.1017/9781108894333

Merkt, M., Lux, S., Hoogerheide, V., Van Gog, T., & Schwan, S. (2020). A change of scenery: Does the setting of an instructional video affect learning? *Journal of Educational Psychology, 112*(6), 1273–1283. https://doi.org/10.1037/edu0000414

Moreno, R., & Mayer, R. E. (2007). Interactive multimodal learning environments. *Educational Psychology Review, 19*, 309–326. https://doi.org/10.1007/s10648-007-9047-2

Murnane, K., Phelps, M. P., & Malmberg, K. (1999). Context-dependent recognition memory: The ICE theory. *Journal of Experimental Psychology: General, 128*(4), 403–415. https://doi.org/10.1037/0096-3445.128.4.403

Parker, A., Ngu, H., & Cassaday, H. J. (2001). Odour and proustian memory: Reduction of context-dependent forgetting and multiple forms of memory. *Applied Cognitive Psychology, 15*(2), 159–171. https://doi.org/10.1002/1099-0720(200103/04)15:2%3c159::AID-ACP694%3e3.0.CO;2-D

Peng, X., Xu, Q., Chen, Y., Zhou, C., Ge, Y., & Li, N. (2021). An eye tracking study: Positive emotional interface design facilitates learning outcomes in multimedia learning? *International Journal of Educational Technology in Higher Education, 18*(1), 40. https://doi.org/10.1186/s41239-021-00274-x

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. https://doi.org/10.1016/j.learninstruc.2013.02.006

Pointer, S. C., & Bond, N. W. (1998). Context-dependent memory: Colour versus odour. *Chemical Senses, 23*(3), 359–362. https://doi.org/10.1093/chemse/23.3.359

Prestera, G., Clariana, R., & Peck, A. (2005). Memory-context effects of screen color in multiple-choice and fill-in tests. *Journal of Educational Multimedia and Hypermedia, 14*(4), 415–436.

Rutherford, A. (2004). Environmental context-dependent recognition memory effects: An examination of ICE model and cue-overload hypotheses. *The Quarterly Journal of Experimental Psychology Section A, 57*(1), 107–127. https://doi.org/10.1080/02724980343000152

Sahlins, M. (1976). Colors and cultures. *Semiotica, 16*(1), 1–22. https://doi.org/10.1515/semi.1976.16.1.1

Sakai, T., Isarida, T. K., & Isarida, T. (2010). Context-dependent effects of background colour in free recall with spatially grouped words. *Memory, 18*(7), 743–753. https://doi.org/10.1080/09658211.2010.508748

Sattarzadeh, M. H., & Tahmasebi Boroujeni, S. (2021). Does red background color enhance short-term memory or blue? *International Journal of Motor Control and Learning, 3*(3), 57–66. https://doi.org/10.52547/ijmcl.3.3.57

Schneider, S., Beege, M., Nebel, S., Schnaubert, L., & Rey, G. D. (2022). The cognitive-affective-social theory of learning in digital environments (CASTLE). *Educational Psychology Review, 34*(1), 1–38. https://doi.org/10.1007/s10648-021-09626-5

Shangguan, C., Wang, Z., Gong, S., Guo, Y., & Xu, S. (2020). More attractive or more interactive? The effects of multi-leveled emotional design on middle school students' multimedia learning. *Frontiers in Psychology, 10*, 3065. https://doi.org/10.3389/fpsyg.2019.03065

Skulmowski, A. (2022). When color coding backfires: A guidance reversal effect when learning with realistic visualizations. *Education and Information Technologies, 27*(4), 4621–4636. https://doi.org/10.1007/s10639-021-10796-6

Skulmowski, A., & Rey, G. D. (2021). Realism as a retrieval cue: Evidence for concreteness-specific effects of realistic, schematic, and verbal components of visualizations on learning and testing. *Human Behavior and Emerging Technologies, 3*(2), 283–295. https://doi.org/10.1002/hbe2.209

Smith, S. M., & Vela, E. (2001). Environmental context-dependent memory: A review and meta-analysis. *Psychonomic Bulletin & Review, 8*(2), 203–220. https://doi.org/10.3758/BF03196157

Smith, S. M., & Manzano, I. (2010). Video context-dependent recall. *Behavior Research Methods, 42*(1), 292–301. https://doi.org/10.3758/BRM.42.1.292

Sun, C.-Y., & Liu, D.-Z. (2016). Effects of the background color of learning materials on cognitive load and learning performance. *Journal of Psychological Science, 39*(4), 869–874.

Sweller, J., Van Merriënboer, J. J. G., & Paas, F. G. W. C. (1998). Cognitive architecture and instructional design. *Educational Psychology Review, 10*(3), 251–296. https://doi.org/10.1023/A:1022193728205

Sweller, J., Van Merriënboer, J. J. G., & Paas, F. (2019). Cognitive architecture and instructional design: 20 years later. *Educational Psychology Review, 31*(2), 261–292. https://doi.org/10.1007/s10648-019-09465-5

Tulving, E., & Pearlstone, Z. (1966). Availability versus accessibility of information in memory for words. *Journal of Verbal Learning and Verbal Behavior, 5*(4), 381–391. https://doi.org/10.1016/S0022-5371(66)80048-8

Uzun, A. M., & Yıldırım, Z. (2018). Exploring the effect of using different levels of emotional design features in multimedia science learning. *Computers & Education, 119*, 112–128. https://doi.org/10.1016/j.compedu.2018.01.002

Wang, X., Mayer, R. E., Han, M., & Zhang, L. (2023). Two emotional design features are more effective than one in multimedia learning. *Journal of Educational Computing Research, 60*(8), 1991–2014. https://doi.org/10.1177/07356331221090845

Wierzbicka, A. (1990). The meaning of color terms: Semantics, culture, and cognition. *Cognitive Linguistics, 1*(1), 99–150. https://doi.org/10.1515/cogl.1990.1.1.99

Wu, X.-Y. (2023). Exploring the effects of digital technology on deep learning: A meta-analysis. *Education and Information Technologies*. https://doi.org/10.1007/s10639-023-12307-1

Wong, R. M., & Adesope, O. O. (2021). Meta-analysis of emotional designs in multimedia learning: A replication and extension study. *Educational Psychology Review, 33*(2), 357–385. https://doi.org/10.1007/s10648-020-09545-x

Zhou, Z. (2018). Second language learning in the technology-mediated environments. *Asian Education Studies, 3*(1), 18. https://doi.org/10.20849/aes.v3i1.307

Zorko, A., Ivančić Valenko, S., Tomiša, M., Keček, D., & Čerepinko, D. (2017). The impact of the text and background color on the screen reading experience. *Tehnički Glasnik, 11*(3), 78–82.

Authors and Affiliations

Felicia Meusel[1] · Nadine Scheller[1] · Günter Daniel Rey[1] · Sascha Schneider[2]

✉ Felicia Meusel
felicia.meusel@phil.tu-chemnitz.de

Nadine Scheller
nadine.scheller@phil.tu-chemnitz.de

Günter Daniel Rey
guenter-daniel.rey@phil.tu-chemnitz.de

Sascha Schneider
sascha.schneider@ife.uzh.ch

1 Psychology of Learning With Digital Media, Institute for Media Research, Faculty of Humanities, Chemnitz University of Technology, Straße der Nationen 12, 09111 Chemnitz, Germany

2 Educational Technology, Institute of Education, University of Zurich, Zurich, Switzerland