frontiers | Frontiers in Education

TYPE Original Research
PUBLISHED 15 June 2023
DOI 10.3389/feduc.2023.1161460

OPEN ACCESS

EDITED BY
Antonio Palacios-Rodríguez,
Sevilla University, Spain

REVIEWED BY
Mona Wong,
Yew Chung College of Early Childhood Education, Hong Kong SAR, China
Aslina Baharum,
MARA University of Technology, Malaysia

*CORRESPONDENCE
Maik Beege
✉ maik.beege@ph-freiburg.de

RECEIVED 08 February 2023
ACCEPTED 23 May 2023
PUBLISHED 15 June 2023

CITATION
Beege M, Scherer D and Weiß E (2023) Improving e-learning websites: the role of menu depth and metacognitive support. *Front. Educ.* 8:1161460.
doi: 10.3389/feduc.2023.1161460

Improving e-learning websites: the role of menu depth and metacognitive support

Maik Beege[1]*, Demian Scherer[2] and Elena Weiß[3]

[1]Digital Media in Education, Department of Psychology, University of Education, Freiburg, Germany, [2]Department for Psychology in Education, University of Münster, Münster, Germany, [3]Psychology of Learning With Digital Media, Department for Media Research, Chemnitz University of Technology, Chemnitz, Germany

Introduction: Results from experimental research in instructional psychology imply that a deep menu structure of a e-learning website may provide useful segmentation. However, menu depth also increases the need for navigation and thus, might have impairing effects on learning. Furthermore, instructional support can be provided by including a checklist, to ensure that learners reflect on their study progress. The study aimed at investigating which menu structure is beneficial for e-learning websites and whether a checklist could compensate the negative effects of an unfavorable menu structure.

Methods: Therefore, in an online experiment, we let 101 students learn facts about rocks from an e-learning website with either a deep or a flat menu structure. We further manipulated whether metacognitive support through a checklist was provided or not. Learning outcomes, cognitive load, metacognitive factors as well as learning time were measured.

Results: Results show no main effects of the menu depth or the presence of a checklist on retention and transfer performance. Learning achievements in percent for retention were 37.31 (deep menu/checklist), 31.10 (deep menu/no checklist), 36.07 (flat menu/checklist), 38.13 (flat menu, no checklist) and for transfer were 35.19 (deep menu/checklist), 34.40 (deep menu/no checklist), 37.78 (flat menu/checklist), 33.23 (flat menu, no checklist). Yet, there are hints that the deeper menu structure had a negative effect on learning processes: The deep menu structure led to an enhanced extraneous cognitive load (ECL) and reduced learning efficiency. However, providing a checklist had beneficial effects mainly when learning with a deep menu structure but not overall. Unexpectedly, the presence of the checklist did not influence metacognitive measures.

Discussion: Our study suggests that possible costs of a deep menu structure should be considered when designing instructional checklists. However, the study also provides a way in which these costs can be compensated, which is by using a checklist. Implications for instructional research and e-learning are discussed.

KEYWORDS

e-learning, website design, menu structure, metacognitive support, checklist

1. Introduction

E-Learning websites become increasingly important in educational scenarios in tertiary and secondary education (e.g., creating a website for information distribution or course enrollment). Technological innovations made even complex multimedia websites available for everyone and thus Web-based instruction got in the focus of early (e.g., Lee, 2001) and actual research (Toan et al., 2021). As e-learning websites quickly received considerable attention as a means of providing alternatives to traditional face-to-face, instructor-led education (Douglas and Van Der Vyver, 2004), calls have been made for research into an effective design that promotes learning. For example, Cook and Dupras (2004) outlined that active learning principles (e.g., a clear and consistent navigation) improve e-learning websites. Considering online courses, pedagogical usability (usability as this affects educational website design and development) was brought to the fore and is researched until today (Pham et al., 2021). In particular during the COVID-19 pandemic, e-learning websites became important for information sharing or explicit instruction. Consequently, the current study investigates a popular e-learning portal in Germany (OPAL; BPS Bildungsportal Sachsen GmbH, 2022) with the aim to increase instructional quality. In line with recent research, the menu structure (Prezenski and Russwinkel, 2014) as well as metacognitive support through a checklist (Ukrayinska, 2020) were the focus of research.

1.1. Learning with multimedia websites

The Cognitive Theory of Multimedia Learning (CTML; Mayer, 2021) is a prominent theory that addresses learning with multimedia. The theory is based on four fundamental assumptions. First, it posits that the human information-processing system consists of two channels: a visual/pictorial channel and an auditory/verbal channel. Second, each channel has a limited capacity. Third, the theory postulates three memory systems: sensory memory, working memory, and long-term memory. Finally, learning is viewed as an active process involving coordinated cognitive processes, which can be further defined as selecting, organizing, and integrating (Mayer, 2021). Selecting information refers to the act of paying attention to spoken or written words presented on a website, as they pass through the auditory sensory memory. The learner creates a mental representation of selected words or phrases in their verbal working memory, while selected pictures are represented in their visual working memory. Organizing involves establishing connections between pieces of knowledge to form a coherent mental model of verbal and visual information separately. Finally, integrating entails connecting the verbal and pictorial models with each other, as well as with the learner's prior knowledge stored in long-term memory.

During active processing, the working memory is loaded. Here, a second theory, the cognitive load theory (CLT; Sweller et al., 2019; Sweller, 2020), becomes relevant. The theory distinguishes between two types of cognitive load during the learning process. First, there is the productive cognitive load, which arises from processes directly involved in learning (Kalyuga and Singh, 2016). This productive load encompasses the intrinsic cognitive load (ICL), which is influenced by the complexity of the information being learned (i.e., element interactivity), domain-specific prior knowledge (Sweller et al., 2019), and germane resources (GR; Krieglstein et al., 2022). Germane resources refer to the mental effort invested in dealing with the ICL and constructing schemas (Paas and van Merriënboer, 2020). Second, there is the unproductive or extraneous cognitive load (ECL), which arises from processes that are not relevant for and may even hinder learning (Kalyuga and Singh, 2016). ECL arises from the presentation and design of the website (Sweller et al., 2019). When information is presented in an unfavorable manner, it leads to unnecessary utilization of cognitive resources that could otherwise be allocated to learning-relevant processes.

Both theories are relevant considering learning with websites. The CTML emphasizes the importance of the visual channel with limited resources for processing the information presented on the website. The CLT further specifies these processes by determining whether learning-relevant or learning-irrelevant processes are triggered.

1.1.1. Menu structure

When discussing the effect of a website's menu structure on learning, two central strands of research have to be considered: (1) the segmenting effect and (2) research regarding website navigation. The segmenting effect posits that individuals exhibit improved learning outcomes when multimedia instructions are presented in meaningful and coherent segments rather than as continuous units (Mayer and Pilegard, 2014). The segmenting effect can be explained by three theoretical explanations, which are not mutually exclusive (Spanjers et al., 2010). Firstly, segmenting enhances learning by reducing the cognitive load associated with chunking and structuring the e-learning website, as these processes have already been performed by the designer. Consequently, learners who receive multimedia instructions presented as continuous units may encounter difficulties in organizing and structuring the instruction into meaningful segments. This facilitates the selecting and organizing processes proposed by the Cognitive Theory of Multimedia Learning (CTML; Mayer, 2021) and, as a result, enhances productive load (Kalyuga and Singh, 2016). Second, segmenting is effective because it allows learners more time to process the multimedia instruction, specifically the information presented on different sub-websites. Learners who receive all the information on a single website, without the opportunity to pause and reflect on the learning process after reviewing sub-websites, may experience cognitive overload at certain points during the instruction, exceeding their working memory capacity for retaining information (Wickens et al., 2013). Conversely, learners who are provided with a segmented e-learning website have sufficient time for cognitive processing of the information on sub-websites and ample opportunity to mentally rehearse the multimedia instruction, thereby avoiding cognitive overload (Schnotz and Lowe, 2008; Stiller et al., 2011). Third, segmenting enables learners to adapt the presentation pace to their individual needs (e.g., Hasler et al., 2007). By organizing coherent information into sub-websites, learners can reflect on each coherent unit for as long as necessary. This perception of control over the task may lead to higher learning performance (Wouters, 2007), based on the aforementioned explanations. A recent meta-analysis (Rey et al., 2019) outlined that segmentation reduces the overall cognitive load and increases learning time and learning performance. Segmentation is effective considering system-paced as well as learner-paced learning materials like menu-based websites.

Research regarding websites offers a deeper look into learning processes when using e-learning websites with different menu structures. In line with the segmenting principle and the CTML, a deep (and thus, more segmented) menu structure decreases information load (Snowberry et al., 1983; Kurtenbach, 1993) and increases time spent on the website (Prezenski and Russwinkel, 2014). Nevertheless, studies revealed that this additional time is not necessarily productive. As deepening the menu structure is not just segmenting information but also increasing the need for navigation through the website, a deeper, more segmented website leads to misnavigation and an unnecessary increase in search time (Snowberry et al., 1983; Samp, 2013; Prezenski and Russwinkel, 2014). Consequently, unproductive load increases (Kalyuga and Singh, 2016). In particular, novices who have not worked with a certain website before, struggle with deeper menus (Cockburn et al., 2007). These results are further reflected in studies that investigated subjective evaluations of website users (e.g., Geven et al., 2006). The authors emphasized that regardless of the technical device, users prefer narrow menu hierarchies. Studies with regard to menu structure in explicit learning contexts are rare. For example, Patsula et al. (2010; see further: Farris et al., 2002) investigated menu structure with regard to multiple variables and concluded that a structured menu led to enhanced retention performance and fewer navigational errors. In a second study, the benefits of a structured menu with regard to retention performance were larger than one standard deviation (Patsula et al., 2010). Summarizing, deepening the menu structure might have beneficial effects, considering the segmenting effect (e.g., Rey et al., 2019) but because of the additional need for navigation, negative effects on learning outcomes are possible as well (Patsula et al., 2010). However, an unfavorable menu structure can possibly be compensated for by integrating additional support on the e-learning website. This study aimed to investigate whether a checklist can act as metacognitive support that might compensate for suppressing the effects of deep or flat menu structures.

1.1.2. Metacognitive support

Metacognition describes the awareness and regulation of one's cognitive processes while learning (cognition of cognition; Flavell, 1979). Basically, it is assumed that the ability to monitor, regulate, and control one's own cognitive resources is beneficial for learning (Akturk and Sahin, 2011). This study focuses on the ability of metacognitive monitoring. Monitoring means that the current state of learning is assessed by the learner based on how well the material has already been understood (Thiede et al., 2003). It therefore seems logical that learning with e-learning websites also requires the ability to monitor one's own learning progress (e.g., Mudrick et al., 2019). The theoretical basis of monitoring processes during media-based learning can be derived from the Cognitive-Affective Theory of Learning with Media (CATLM; Moreno and Mayer, 2007), which has been proposed as a further development of the CTML. The CATLM assumes that learners may use metacognitive skills to regulate their cognitive processing needed for understanding. Accordingly, metacognitive factors may have a mediating role while learning (e.g., McGuinness, 1990).

Whereas, multiple studies investigate the beneficial role of so-called metacognitive prompts during learning (e.g., Zheng, 2016; Guo, 2022), the current study examined a design feature, which can be easily implemented in e-learning websites: checklists. The goal of a checklist is to sensitize learners to identify information units which have not yet been fully learned and to encourage them to navigate to these units. Learners can reflect if they have reviewed all relevant sub-websites or if they have missed some information due to a suppressing menu structure. Consequently, learners can reflect on the processing of ICL and can regulate if additional effort (the additional investment of GR) is necessary. According to Rowlands (2007), well-designed checklists identify steps that students can take to complete complex learning tasks, which scaffolds students' metacognitive development and fosters confidence and independence in the task. Including completion checklists lead to a high level of involvement in cooperation scenarios (Corpuz-Abenoja, 2022), but even in individual scenarios, checklists are beneficial (Ukrayinska, 2020). In particular because of the comparative ease of development and implementation in learning scenarios, checklists are a multifunctional tool for developing students' assessment literacy as this process involves both learning and evaluating (Ukrayinska, 2020).

Empirical research on metacognitive monitoring often concentrates to examine whether people are able to accurately predict their learning performance as the accurate assessment of one's own learning progress is a crucial predictor of successful learning (Son and Metcalfe, 2005). As an indicator for metacognitive monitoring, this accuracy between monitoring and performance can be measured with judgment of learning scales (JOL; Dunlosky and Thiede, 2013). For example, Beege et al. (2021) asked learners to think about how well they will perform in a learning test that will deal with the previously learnt information. Another prominent measure is the retrospective confidence (RC; Dinsmore and Parkinson, 2013). RC assesses the confidence of the performance in a learning test. When learners are confronted with learning websites with a menu structure that might suppress learning, metacognitive support may be particularly helpful. When learners have difficulties to navigate through relevant information, an additional checklist might strengthen the learner's ability to self-regulate their learning process.

1.2. Hypotheses

Regarding the menu structure, a deeper menu structure might enhance learning because of the segmenting effect (Rey et al., 2019). Nevertheless, suppressing effects are possible as well (Patsula et al., 2010). Consequently, either a flat menu structure enhances GLC and reduces ECL or a deep menu structure enhances productive load and reduces unproductive load in contrast to its counterpart. The following contrary hypotheses were formulated.

H1a: Learners receiving an e-learning website with a deep menu structure outperform learners receiving an e-learning website with a flat menu structure.

H1b: Learners receiving an e-learning website with a flat menu structure outperform learners receiving an e-learning website with a deep menu structure.

H2a: Learners receiving an e-learning website with a deep menu structure report reduced unproductive load and increased productive load in contrast to learners receiving an e-learning website with a flat menu structure.

H2b: Learners receiving an e-learning website with a flat menu structure report reduced unproductive load and increased productive load in contrast to learners receiving an e-learning website with a deep menu structure.

According to research regarding the segmenting effect (Tabbers and de Koeijer, 2010) and menu structures (Prezenski and Russwinkel, 2014), a deep menu structure should increase learning time.

H3: Learners receiving an e-learning website with a deep menu structure have a longer learning time than learners receiving an e-learning website with a flat menu structure.

Results regarding the metacognitive support of the checklist highly depend on the potential effect of the menu structure on learning-relevant variables. It is hypothesized that metacognitive support is particularly beneficial when the design of the menu structure is suppressive for learning. Nevertheless, based on a recent meta-analysis (Guo, 2022), a checklist should generally be beneficial for learning. This should be reflected in learning outcomes as well as productive load scores.

H4: Learners receiving an e-learning website with a checklist outperform learners receiving an e-learning website without a checklist.

H5: Learners receiving an e-learning website with a checklist report increased productive load in contrast to learners receiving an e-learning website without a checklist.

H6: Interaction Hypothesis 1: A checklist is particularly effective for learning outcomes and productive load when learners receive an e-learning website with a suppressing menu structure.

Regarding the checklist, metacognitive variables are important to consider. Metacognitive support through a checklist should foster monitoring processes (Rowlands, 2007). This should be reflected in enhanced JOL and RC scores as well as in enhanced metacognitive accuracy. Again, metacognitive benefits should be particularly relevant and beneficial when the design of the menu structure is suppressive for learning.

H7: Learners receiving an e-learning website with a checklist report increased metacognitive judgments and reach an increased metacognitive accuracy in contrast to learners receiving an e-learning website without a checklist.

H8: Interaction Hypothesis 2: A checklist particularly increases metacognitive judgments and accuracy when learners receive an e-learning website with a suppressing menu structure.

Additional variables were explored to get a deeper insight into the learning process. As learning time was measured, instructional efficiency was investigated. Furthermore, the subjective effectiveness of the checklist was measured.

RQ1: Do menu depth and a checklist influence instructional efficiency?

RQ2: Does the subjective effectiveness of the checklist differ with respect to the menu structure?

2. Methods

2.1. Participants and design

A recent meta-analysis with regard to learner-paced segmenting effect found a medium effect with regard to transfer performance ($d = 0.45$; Rey et al., 2019). However, as segmenting was operationalized through menu structure, additional research was reviewed for power analysis. Patsula et al. (2010) outlined that structured menu designs increased the mean value of retention performance with over one standard deviation compared to the mean value of unstructured menu designs. A deep menu structure increased false navigation with a high effect size (Snowberry et al., 1983) and increased search time with a particular high effect size ($\eta^2 = 0.343$; Prezenski and Russwinkel, 2014) in contrast to a flat menu structure. Studies with regard to metacognitive prompts revealed a medium effect size with regard to self-regulated learning as well ($d = 0.50$; Guo, 2022). Consequently, the current study was powered for a medium to large effect size to consider effects regarding segmenting and menu structure (GPower, mean of $f = 0.25$ and $f = 0.40$; Erdfelder et al., 1996). According to an a priori power analysis ($f = 0.325$; $\alpha = 0.05$; $1 - ß = 0.90$; 2×2 design), 102 participants should be acquired. A total of 101 students (75.2% female; age: $M = 22.22$; $SD = 2.80$) from Chemnitz University of Technology (62.4%) and Freiburg University of Education (37.6%) participated in this experiment. Participants were university students (98%) or employed (2%).

The experiment was carried out as a 2×2 design. Each student was randomly assigned to one cell of a between-subjects design by an automatic randomization system (menu structure: flat menu vs. deep menu and metacognitive support: checklist vs. no checklist). Twenty-two students were assigned to the condition with the flat menu structure and without a checklist, 23 students were assigned to the condition with the flat menu structure and with a checklist, 29 students were assigned to the condition with the deep menu structure and without a checklist, and 27 students were assigned to the condition with the deep menu structure and with a checklist.

No significant differences with regard to the between-subject factors existed in terms of age, prior knowledge or prior experience with OPAL, $F_{(3,97)} = (0.11; 0.71)$; $p = (0.55; 0.95)$, and university or gender, $\chi 2 = (2.35; 6.21)$; $p = (0.40; 0.50)$.

2.2. Materials

The learning material consisted of a self-created e-learning website on the learning platform OPAL (BPS Bildungsportal Sachsen GmbH, 2022). The website was structured like an actual geology university course. First, information about the preliminary examination and examination performance was presented. Afterward, the website included facts about properties, formation, and use of rocks. This content was chosen because prior knowledge is generally considered to be low among most populations of students. Furthermore, the information could be easily divided into subtopics. Consequently, any type of rock was presented in a separate chapter. The chapters included an instructional text about the rocks and additional pictures for visualization. On the left side, a menu was displayed that allowed the learner to navigate through the chapters. In the conditions with a checklist, the checklist was additionally displayed in the menu. A screen example of the website is displayed in Figure 1.

2.2.1. Menu structure

With respect to literature regarding the menu structure of websites (e.g., Prezenski and Russwinkel, 2014), the depth of the navigation and information presentation was varied. In the flat menu structure condition, there was no need for additional navigation through a menu. All information chapters were presented on the homepage. Yet, the division into chapters was clearly visible (see Figure 2). In the deep menu structure condition, there was one additional structure level. Learners had to click on the sub-topics to get to the information within the chapter. Consequently, the amount of information presented at the same time was varied. In the flat menu structure condition, there was no need for additional navigation but a potential overload due to the amount of simultaneously presented information. In contrast, in the deep menu structure condition, the amount of simultaneously presented information was reduced, but additional navigation was necessary to be able to retrieve information.

2.2.2. Metacognitive Support

Figure 3 illustrates that the checklist was implemented within the menu structure on the left side of the website (see Figure 3). The checklist mapped all the chapters with their respective subtopics. Thus, it was ensured that, despite the overload, learners in the flat menu structure condition could reflect that they had read all the information. In the deep menu structure, learners could reflect that they had navigated to all information. At the beginning of the learning phase, all bullet points are listed under "open". As soon as learners check off a bullet point of the checklist, it will be listed further down in the "done" section. Learners could open and work on the checklist at any time during the learning phase. Learners could further click on bullet points in the "done" section and put them back to the "open" section. Consequently, learners could reflect how far they have already progressed and reflect to previously learned chapters at any time.

2.3. Measures

2.3.1. Metacognitive judgments

The methodology employed to measure metacognitive judgments was based on the work of Pieger et al. (2016). The Judgment of Learning (JOL) was assessed following the learning phase to determine the learners' estimations of their ability to answer questions related to the material they had just studied. Participants were asked to indicate the percentage of questions they believed they could answer correctly, using a scale ranging from 0 (no questions) to 100 (all questions). Retrospective confidence (RC) was measured by asking participants to estimate the percentage of study questions they believed they had answered correctly. This question was rated on a scale from 0 (no correct answers) to 100 (all answers correct). The RC question was administered after the learning scales had been completed.

2.3.2. Checklist rating

A self-created questionnaire was implemented in both conditions with a checklist to assess the use and subjective usefulness of the list for the learning process. A 6-item questionnaire ($\alpha = 0.75$) was implemented after the JOL question, and participants had to rate the items (e.g., "I have actively used the checklist.") on a 7-point Likert scale ranging from 1 (does not apply at all) to 7 (fully applies).

2.3.3. Cognitive load

Cognitive load was assessed using the self-reported scale developed by Klepsch et al. (2017), which was selected due to its relevance in measuring the complexity of the content and the recognition of important information. Intrinsic load (ICL) was measured by two items ($\alpha = 0.80$), such as "This task was very complex." Extraneous load (ECL) was assessed using three items ($\alpha = 0.83$), including "During this task, it was exhausting to find the important information." Germane resources (GR) were measured with three items ($\alpha = 0.66$), such as "My goal while working on the task was to understand everything correctly." Participants rated these items on a 7-point Likert scale, ranging from 1 (absolutely wrong) to 7 (absolutely correct).

2.3.4. Learning time

As studies regarding menu structure revealed important effects with regard to navigation time (Prezenski and Russwinkel, 2014), learning time was assessed. Therefore, the time that participants worked with the e-learning website was tracked.

2.3.5. Knowledge measures

First, prior knowledge was measured with four open-answer questions about the definition, types, emergence, and characteristics of rocks ($\alpha = 0.82$). Students were asked to write as many facts as they know. They gained a point for each information, which was part of the learning material. Students were told to write "I don't know" if they did not know any information. The inter-rater reliability of two pre-trained reviewers was high, ICC $(1, k) = (0.86, 0.98)$, $F_{(100, 100)} = (13.42, 106.84)$, $p < 0.001$ (Koo and Li, 2016). Overall, prior knowledge was low (mean points: $M = 2.31$; $SD = 2.11$ with a maximum of 11 points).

Second, retention was measured with a 12-item questionnaire ($\alpha = 0.65$). Four multiple-choice questions (e.g., "Which statements about sedimentary rocks are correct?") were included. Students were asked to choose between up to four possible answers; one up to all answer options could be correct. Students received points for selecting correct answers and for not selecting incorrect answers. Furthermore, eight open-answer questions (e.g., "Into what types can sedimentary rocks be divided?") were included. The inter-rater reliability was high, ICC $(1, k) = (0.91, 0.99)$, $F_{(100, 100)} = (20.75, 158.11)$, $p < 0.001$. Retention questions covered information that was explicitly presented within the learning material. Students were able to get up to 27 points.

10.3389/feduc.2023.1161460

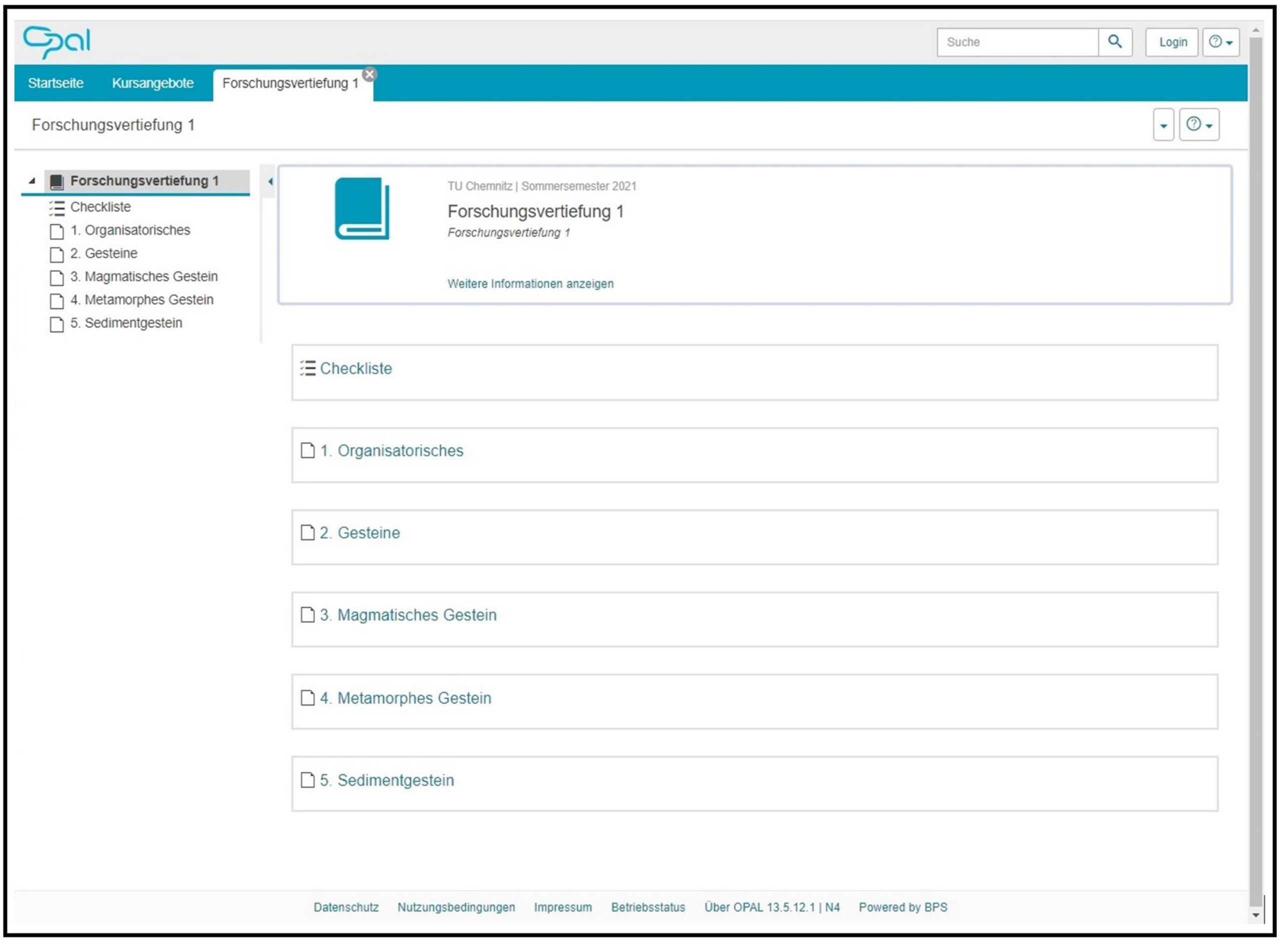

FIGURE 1
Screen example of the website (deep menu structure and with a checklist).

Whereas, retention can be defined as remembering or reproducing information, which was presented in the instructional text, transfer refers to applying the knowledge in order to solve novel problems, which were not explicitly presented in the learning material (Mayer, 2014). A 9-item scale with open-answer questions ($\alpha = 0.79$; e.g., "Which type of rock is the costliest to procure by man? Explain why!") was created in which every item presented a new scenario. Every scenario could be solved with the knowledge the students had obtained from the learning material. The inter-rater reliability was moderate-to-high ICC, $(1, k) = (0.94, 0.97)$, $F_{(100, 100)} = (35.04, 64.90)$, $p < 0.001$ or perfect. Students could reach up to 21 points.

Finally, learning efficiency as well as measures of absolute accuracy of metacognitive judgments and corresponding bias measures were investigated. To determine the absolute accuracy of one's confidence judgments, the difference between the estimated proportion of correct responses (percentage divided by 100) and the proportion of correct responses (for both retention and transfer items) was calculated and squared (Schraw, 2009). To calculate the bias measure the same difference was used but not squared. We used the proportion of correct responses based on both retention and transfer performance, as participants also gave an overall estimate about the percentage of their correct responses. As we have two estimates for the proportion of correct responses, one obtained during and one after reading, two such indices for the absolute accuracy were calculated.

$$Accuracy_{JOL} = (proportion_{JOL} - proportion_{correct})^2$$

$$Accuracy_{RC} = (proportion_{RC} - proportion_{correct})^2$$

Please note that these two accuracy indices indicate a discrepancy between a confidence judgment and performance. Therefore, higher values indicate a reduced accuracy.

To calculate efficiency, learning scores (zLretention and zLtransfer) as well as learning time (zT) were z-standardized. Instructional efficiency was calculated with the formula postulated by Van Gog and Paas (2008).

$$Efficiency = \frac{zL - zT}{\sqrt{2}}$$

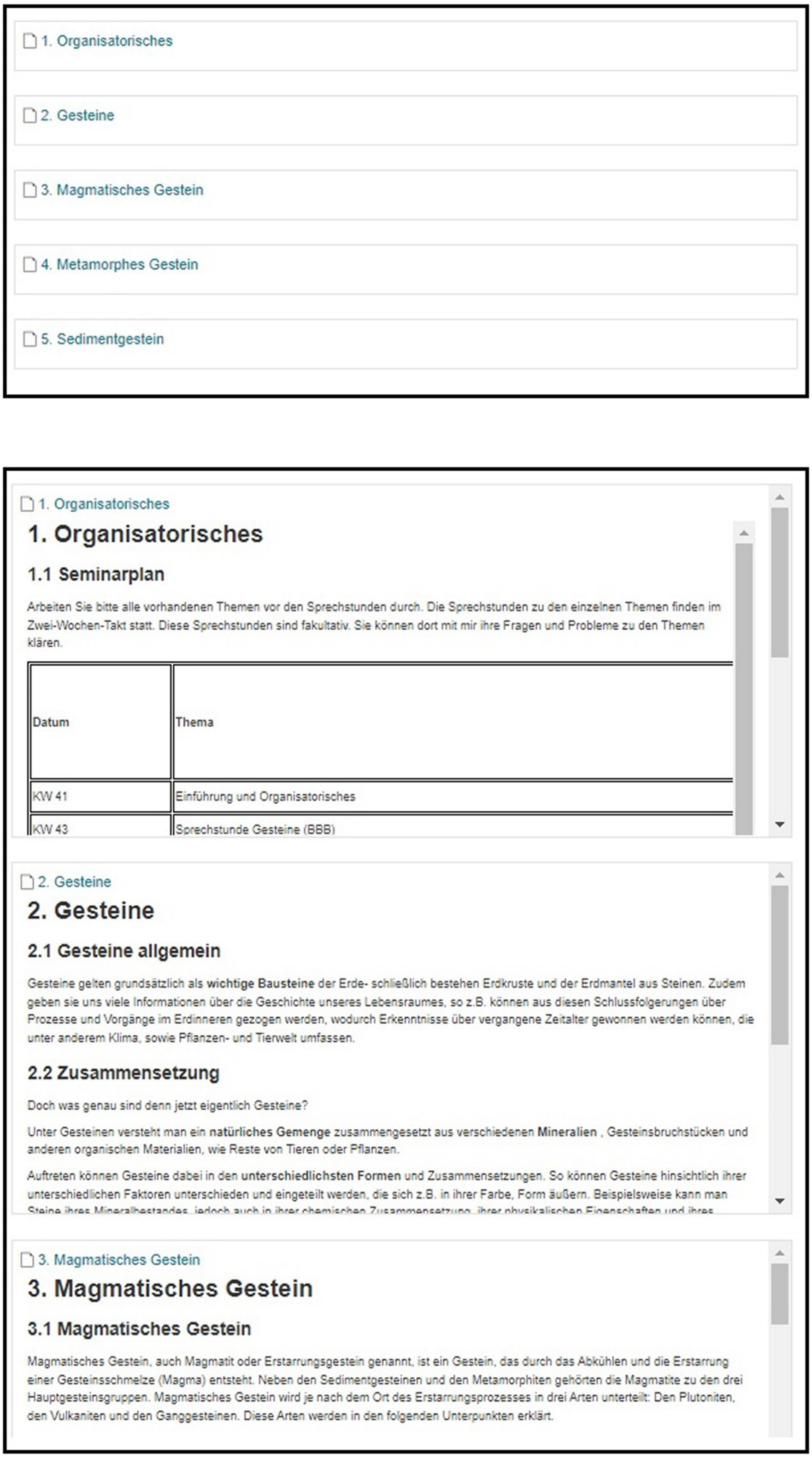

FIGURE 2
Manipulation of menu structure (bottom: flat menu; top: deep menu).

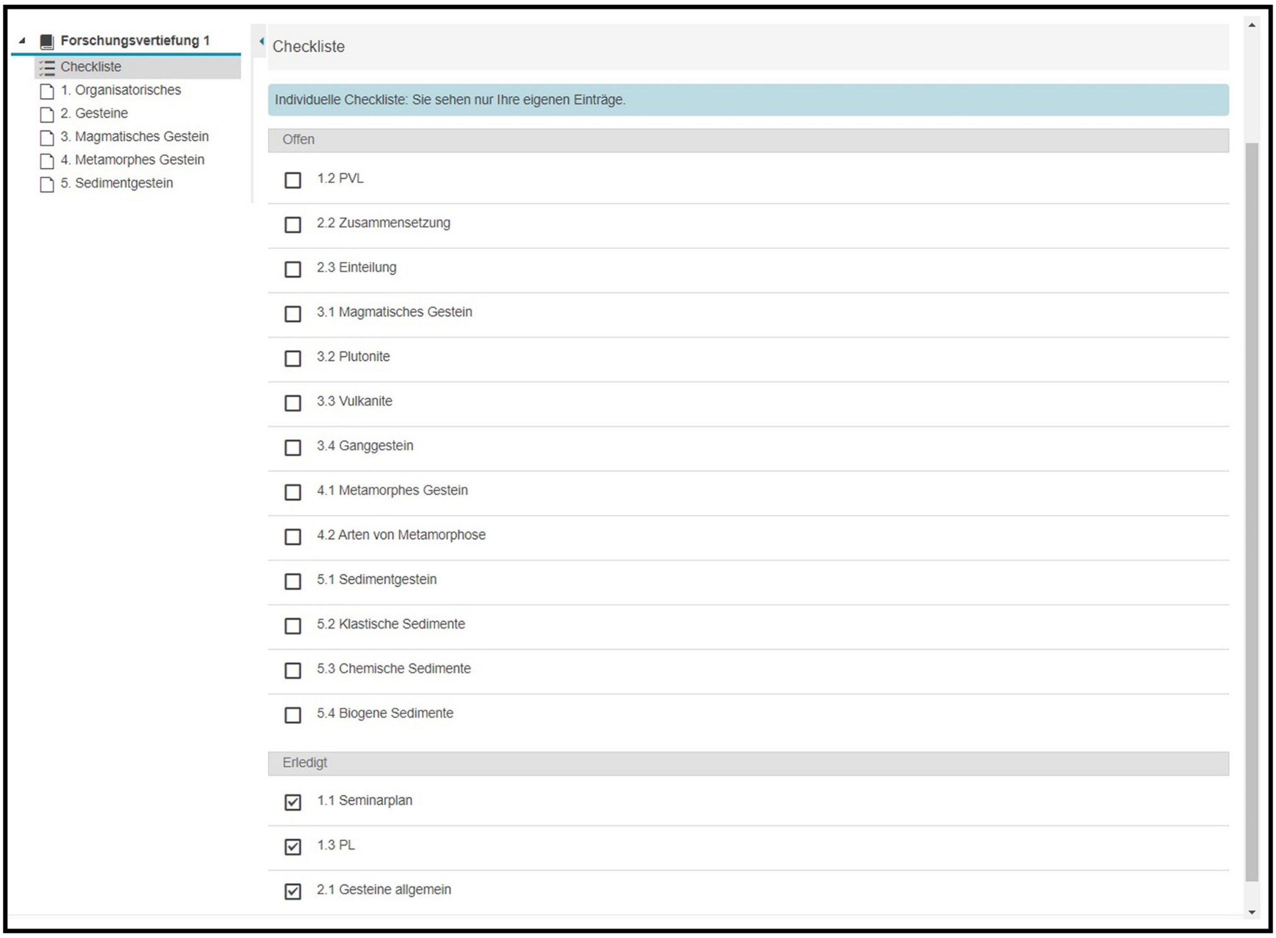

FIGURE 3
Partially edited checklist.

This score was calculated separately for retention and transfer performance as an efficiency regarding learning for retention and as an efficiency with respect to learning for transfer.

2.4. Procedure

The investigation took place in an online environment. Up to four students logged in to the educational platform BigBlueButton. Each student was assigned to a breakout room and received a link for study participation. Students were instructed to share their screens till the end of the learning phase to control participation. Afterward, screen sharing was ended. After receiving the link, the investigation started.

First, the participants were informed that the experiment was a study on the optimization of e-learning websites; they were then asked to answer the prior-knowledge test. Next, they were given the link to the website and were asked to carefully learn all relevant information on the website. Participants could stay on and navigate through the website as long as they wanted. The dependent variables were measured after the learning phase in the following order: JOL, checklist rating (only in the checklist conditions), cognitive load, retention, transfer, RC. Finally, demographic questions were asked. If all tests were completed, the participants could leave the online platform. Altogether, the experiment lasted an average of lasted 45 min.

2.5. Analysis strategy

In the analyses of data, univariate analyses of covariance (ANCOVAs) were conducted in order to assess differences between groups. As prior knowledge significantly correlated with almost every investigated dependent variable,[1] prior knowledge was entered into the analyses of these variables as a covariate. Only

1 With JOL, $r = 0.35$, $p < 0.001$; RC, $r = 0.39$, $p < 0.001$, with ICL $r = -0.21$, $p = 0.034$, with ECL, $r = -0.20$, $p = 0.040$, with GR, $r = -0.25$, $p = 0.013$, retention performance, $r = 0.19$, $p = 0.032$ (one-tailed) 0. For learning efficiency regarding transfer, $r = 0.21$, $p = 0.023$, $\text{Accuracy}_{\text{JOL}}$, $r = 0.23$, $p = 0.020$, $\text{Accuracy}_{\text{RC}}$: $r = 0.173$, $p = 0.083$ (marginally significant), Bias JOL, $r = 0.21$, $p = 0.036$, Bias RC: $r = 0.29$, $p = 0.0030$.

for the invested learning time ($r = 0.03$, $p = 0.799$), the checklist ratings ($r = -0.18$, $p = 0.202$), and the learning efficiency regarding knowledge ($r = 0.11$, $p = 0.276$), there were no substantial correlations with prior knowledge. Accordingly, for these measures, prior knowledge was not entered as a covariate and ANOVAs were used. As the rating of the checklist was only available for the two groups that received a checklist, a t-test instead of an ANOVA was used. Descriptive results for all dependent variables are outlined in Table 1.

3. Results

Based on Table 1, a descriptive analysis revealed that learning scores only slightly differed. Nevertheless, in particular students in the deep menu structure conditions needed significantly more time for the learning phase. Consequently, a deep menu structure reduced instructional efficiency. It further becomes visible that a checklist at least partially compensated the negative efficiency of a deep menu structure in contrast to a flat menu structure. Subjective ratings as well as metacognitive scores only slightly differed, but these results are further investigated using inferential statistical methods.

3.1. Learning outcomes

Regarding knowledge, a 2 × 2 ANCOVA with the factors: menu structure and presence of a checklist and the covariate prior knowledge was performed. This analysis did neither reveal a significant effect of the menu depth manipulation, $F_{(1, 96)} = 2.59$, $p = 0.111$, $\eta_p^2 = 0.026$, nor a significant effect of the checklist, $F_{(1, 96)} = 0.86$, $p = 0.356$, $\eta_p^2 = 0.009$. However, there was a significant interaction of both factors, $F_{(1, 96)} = 5.73$, $p = 0.019$, $\eta_p^2 = 0.059$; that is, for the flat menu structure, there was no significant difference for knowledge items between the condition with a checklist and without a checklist (with numerically higher retention in the condition without a checklist), $F_{(1, 42)}$, $p = 0.211$, $\eta_p^2 = 0.037$, for the effect of the covariate: $F_{(1, 42)} = 4.60$, $p = 0.038$, $\eta_p^2 = 0.099$. In contrast, for the deep menu structure, there was a significant advantage regarding retention performance when learners had a checklist compared to when learners had no checklist, $F_{(1, 53)} = 4.69$, $p = 0.035$, $\eta_p^2 = 0.081$. There was no effect of the covariate, $F_{(1, 53)} = 1.71$, $p = 0.196$, $\eta_p^2 = 0.031$.

The same analysis that is a 2 × 2 ANCOVA with the factors: menu structure and presence of a checklist and the covariate prior knowledge was performed for the transfer performance. This analysis revealed no significant effects besides a significant effect of the covariate prior knowledge, $F_{(1, 96)} = 11.69$, $p < 0.001$, $\eta_p^2 = 0.109$ (all other main effects and interactions: $F < 1$, $p > 0.418$, $\eta_p^2 < 0.007$).[2]

2 For the effect of the checklist, $F_{(1, 96)} = 00.66$, $p = 0.418$, $\eta_p^2 = 0070$. For the effect of menu depth, $F_{(1, 96)} = 00.22$, $p = 0.644$, $\eta_p^2 = 0.0020$. For the interaction, $F_{(1, 96)} = 00.07$, $p = 0.785$, $\eta_p^2 = 0.0010$.

3.2. Checklist rating

In the condition with a deep menu structure, the checklist was rated as significantly more useful ($M = 3.03$, $SD = 1.50$) than in the condition with a flat menu structure ($M = 2.24$, $SD = 0.88$), $t_{(46.47)} = 2.37$, $p = 0.022$, $d = 0.63$ (Levene's test revealed significantly different variances in the two conditions, $F = 5.42$, $p = 0.024$).

3.3. Cognitive load

To check whether there are overall effects on cognitive load, we performed a 2 × 2 MANCOVA with all three cognitive load measures (ICL, ECL, and GR) as dependent variables and with prior knowledge as a covariate. This analysis revealed a significant effect of the factor menu depth, $\Lambda = 0.91$, $F_{(3, 94)} = 3.05$, $p = 0.033$, $\eta_p^2 = 0.089$. The factor checklist closely failed to reach significance, $\Lambda = 0.93$, $F_{(3, 94)} = 2.215$, $p = 0.091$, $\eta_p^2 = 0.066$. There was a significant interaction of both factors, $\Lambda = 0.84$, $F_{(3, 94)} = 6.21$, $p =<.001$, $\eta_p^2 = 0.165$, and a significant effect of the covariate, $\Lambda = 0.84$, $F_{(3, 94)} = 6.14$, $p = < .001$, $\eta_p^2 = 0.164$. To get further insights into effects regarding ICL, ECL, and GR, 2 × 2 analyses of covariance with the factors: menu structure (deep vs. flat) and presence of a checklist (with vs. without) and the covariate prior knowledge were conducted.

Regarding ICL, the 2 × 2 ANCOVA did not reveal a main effect of the factor menu structure, $F_{(1, 96)} = 2.45$, $p = 0.121$, $\eta_p^2 = 0.025$, nor a main effect of the checklist, $F_{(1, 96)} = 0.23$, $p = 0.636$, $\eta_p^2 = 0.002$. However, both factors significantly interacted, $F_{(1, 96)} = 8.39$, $p = 0.005$, $\eta_p^2 = 0.080$; that is, the presence of a list only significantly decreased ICL when the menu structure was deep, $F_{(1, 53)} = 6.07$, $p = 0.017$, $\eta_p^2 = 0.103$J, with $F_{(1, 53)} = 3.56$, $p = 0.065$, $\eta_p^2 = 0.063$ for the effect of the covariate. In contrast, if the menu structure was flat, ICL was numerically (but not significantly) increased by the presence of a checklist, $F_{(1, 42)} = 2.73$, $p = 0.106$, $\eta_p^2 = 0.061$, with $F_{(1, 42)} = 3.39$, $p = 0.073$, $\eta_p^2 = 0.075$. Further, the effect of the covariate prior knowledge was significant, $F_{(1, 96)} = 6.98$, $p = 0.010$, $\eta_p^2 = 0.068$.

For ECL, the corresponding 2 × 2 ANCOVA with the factors menu structure and presence of a checklist revealed a significant main effect of menu structure with overall higher ECL ratings in the conditions with a deep menu structure compared with the conditions with a flat menu structure, $F_{(1, 96)} = 7.62$, $p = 0.007$, $\eta_p^2 = 0.074$. In addition, in the conditions without a checklist ($M = 3.67$, SD $= 1.56$), ECL was higher compared with the conditions with a checklist ($M = 2.85$, SD $= 1.30$) showing that the checklist significantly reduced ECL, $F_{(1, 96)} = 6.54$, $p = 0.012$, $\eta_p^2 = 0.064$. There was a significant effect of the covariate, $F_{(1, 96)} = 7.62$; $p = 0.007$. $\eta_p^2 = 0.074$. Further, the two factors significantly interacted, $F_{(1, 96)} = 11.30$, $p = 0.001$, $\eta_p^2 = 0.105$; that is, the presence of a checklist reduced ECL in the design with a deep menu structure, $F_{(1, 53)} = 19.21$, $p < 0.001$, $\eta_p^2 = 0.266$, with $F_{(1, 53)} = 7.52$, $p = 0.008$, $\eta_p^2 = 0.124$ for the effect of the covariate. However, the presence of a checklist did not influence ELC in the design with a flat menu structure, $F_{(1, 42)} = 0.17$, $p = 0.684$, $\eta_p^2 = 0.004$, with $F_{(1, 42)} = 1.24$, $p = 0.272$, $\eta_p^2 = 0.029$ for the effect of the covariate.

TABLE 1 Means and standard deviations of all dependent variables for the experimental groups.

	Experimental groups							
	Deep menu structure				Flat menu structure			
	Checklist		No checklist		Checklist		No checklist	
	M	SD	*M*	SD	*M*	SD	*M*	SD
Retention (percent)	37.31	12.91	31.10	9.49	36.07	9.02	38.13	6.72
Transfer (percent)	35.19	13.62	34.40	14.32	37.78	14.02	33.23	13.09
ICL	3.52	1.65	4.43	1.34	3.85	1.57	3.30	1.23
ECL	2.80	1.27	4.31	1.55	2.91	1.36	2.83	1.30
GR	5.07	1.32	5.43	1.16	5.78	0.84	5.23	1.18
Learning time (in seconds)	759.88	211.65	966.52	412.00	746.90	240.52	695.27	213.19
JOL	42.33	21.84	47.52	18.67	49.43	20.33	45.59	18.95
RC	38.56	25.98	40.83	23.93	39.74	19.56	39.41	19.97
AC for JOL	0.04	0.04	0.06	0.08	0.05	0.06	0.05	0.06
AC for RC	0.04	0.05	0.06	0.05	0.03	0.04	0.03	0.04
Bias for JOL	0.06	0.18	0.15	0.19	0.13	0.19	0.13	0.199
Bias for RC	0.02	0.21	0.08	0.22	0.03	0.16	0.03	0.16
Learning efficiency (retention)	0.23	0.98	−0.68	1.07	0.17	0.86	0.44	0.77
Learning efficiency (transfer)	0.10	0.87	−0.42	1.38	0.27	0.68	0.15	0.83
Checklist rating	3.38	1.13			2.24	0.88		
Prior knowledge (percent)	20.54	18.71	23.20	18.28	23.32	23.67	16.12	15.92

Please note that for AC (for JOL and RC) higher values indicate an increased discrepancy between confidence judgments and performance. Retention scores ranged from 0 to 27. Transfer scores ranged from 0 to 21. Cognitive Load and Checklist ratings ranged from 1 to 7. JOL and RC ratings ranged from 0 to 100.

For GR, the corresponding 2 × 2 ANCOVA with the factors such as menu structure and presence of a checklist and the covariate prior knowledge revealed a significant effect of the covariate, $F_{(1, 96)} = 5.66$, $p = 0.019$, $\eta_p^2 = 0.056$. There was no significant effect of the menu depth, $F_{(1, 96)} = 1.58$, $p = 0.212$, $\eta_p^2 = 0.016$, and no effect of the presence of a checklist, $F_{(1, 96)} = 0.09$, $p = 0.766$, $\eta_p^2 = 0.001$. The interaction failed to reach significance, $F_{(1, 96)} = 2.91$, $p = .091$, $\eta_p^2 = 0.029$.

3.4. Learning time

Learning time was analyzed with a 2 × 2 ANOVA with the factors: menu structure and presence of a checklist. This analysis revealed a significant main effect of the factor menu structure, with more learning time being invested in the conditions with a deep menu structure, $F_{(1, 97)} = 6.00$, $p = 0.016$, $\eta_p^2 = 0.058$. There was no significant effect of the presence of a checklist, $F_{(1, 97)} = 1.78$, $p = 0.185$, $\eta_p^2 = 0.018$. However, there was a significant interaction of both factors, $F_{(1, 97)} = 4.95$, $p = 0.028$, $\eta_p^2 = 0.049$.

To further investigate this interaction, t-tests were performed, revealing that there was no significant effect of the factor checklist, in the conditions with the flat menu structure with only numerically more learning time being invested when a checklist was available, $t(43) = -0.76$ $p = 0.451$, $d = 0.28$. However, for the deep menu structure, the presence of a checklist significantly decreased learning time, $t_{(54)} = 2.33$, $p = 0.023$, $d = 0.624$.

3.5. Metacognitive judgments

For the analysis of JOL, we conducted a 2 × 2 analysis of covariance with the factors: menu structure (deep vs. flat) and presence of a checklist (with vs. without) and the covariate prior knowledge. For this analysis, there was only a significant effect of the covariate, $F_{(1, 96)} = 13.52$, $p < 0.001$, $\eta_p^2 = 0.123$. All other effects did not reach significance ($F < 1.06$, $p > 0.30$, $\eta_p^2 < 0.011$).

A 2 × 2 ANCOVA was also conducted for the analysis of RC, however, not revealing significant effects (all $F < 0.23$, $p > 0.630$, $\eta_p^2 < 0.003$) besides a significant effect of the covariate, $F_{(1, 96)} = 17.85$, $p < 0.001$, $\eta_p^2 = 0.157$.

3.6. Accuracy scores for JOL and RC

The 2 × 2 ANCOVA of the absolute accuracy score, based on the JOL ratings and the performance score for both knowledge and transfer items, did not reveal significant effects (all $F < 0.62$, $p > 0.433$, $\eta_p^2 < 0.006$) besides a significant effect of the covariate, $F_{(1, 96)} = 5.19$, $p = 0.025$, $\eta_p^2 = 0.051$.

To analyze the absolute accuracy scores based on the RC ratings, a 2 × 2 ANCOVA was performed. This analysis revealed a significant main effect of the factor menu depth, with higher scores (that is reduced metacognitive accuracy) in the conditions with a deep menu structure, $F_{(1, 97)} = 3.96$, $p = 0.049$, $\eta_p^2 = 0.040$. There was no significant effect of the factor checklist, $F_{(1, 96)} = 1.47$,

$p = 0.229$, $\eta_p^2 = 0.015$, and no interaction of both factors, $F_{(1, 96)} = 0.02$, $p = 0.880$, $\eta_p^2 < 0.001$, with $F_{(1, 96)} = 2.82$, $p = 0.096$, $\eta_p^2 = 0.029$, for the effect of the covariate.

To rule out that our manipulations led to biased responses, as a control strategy, we additionally performed corresponding ANCOVAs for the bias measures based on JOL and RC ratings. Neither the ANCOVA for the bias measure based on JOL revealed significant effects (all $F < 1.51$, $p > 0.223$, $\eta_p^2 = 0.016$) besides a significant effect of the covariate, $F_{(1, 96)} = 4.10$, $p = 0.046$, $\eta_p^2 = 0.041$, nor the ANCOVA for the bias for the RC ratings (all $F < 1.09$, $p > 0.230$, $\eta_p^2 = 0.012$) besides again, a significant effect of the covariate, $F_{(1, 96)} = 8.91$, $p = 0.004$, $\eta_p^2 = 0.085$.

3.7. Efficiency scores

The efficiency score for the knowledge items (calculated based on the invested learning time as well as the knowledge scores, see methods) were analyzed with a 2 × 2 ANOVA (factors: checklist and menu structure). There was numerically more efficient learning regarding knowledge acquisition in the conditions with a checklist, $F_{(1, 97)} = 2.95$, $p = 0.089$, $\eta_p^2 = 0.030$. For the factor menu depth, there was a significant main effect with more efficient knowledge acquisition in the condition with a flat menu structure, $F_{(1, 97)} = 7.93$, $p = 0.006$, $\eta_p^2 = 0.076$. There was also a significant interaction of both factors, $F_{(1, 97)} = 9.69$, $p = 0.002$, $\eta_p^2 = 0.091$.

Further investigation of this interaction with t-tests revealed that there was no significant effect of the checklist, for the two conditions with a flat menu structure, with numerically more efficient knowledge acquisition when learning without a checklist, $t_{(43)} = 1.08$, $p = 0.288$, $d = 0.32$. However, for the two conditions with a deep menu structure, learning with a checklist significantly improved the efficient knowledge acquisition, $t_{(54)} = 3.31$, $p = 0.002$, $d = 0.89$.

To investigate the learning efficiency to solve transfer items, a 2 × 2 ANCOVA was performed (with the factors menu depth and checklist and prior knowledge as a covariate). There was only numerically higher efficiency in the conditions learning with a checklist, $F_{(1, 96)} = 2.18$, $p = 0.144$, $\eta_p^2 = 0.022$. There was, however, significantly more efficient learning for solving transfer items in the condition with a flat menu structure, $F_{(1, 96)} = 4.02$, $p = 0.048$, $\eta_p^2 = 0.040$. The two factors did not interact, $F_{(1, 96)} = 1.78$, $p = 0.185$, $\eta_p^2 = 0.018$. There was a significant effect of the covariate, $F_{(1, 96)} = 5.87$, $p = 0.017$, $\eta_p^2 = 0.058$.

4. Discussion

The study aimed at investigating the effects of the menu structure and checklists and learning with e-learning websites. A deep menu structure neither fostered nor inhibited learning outcomes. Consequently, H1a as well as H1b could not be supported. A deep menu structure increased ECL but there was no main effect regarding the menu structure for ICL and GR. Thus, H2b could be partially supported and H2a could not be supported. In line with H3, learners who received a website with a deep menu structure learned longer than learners who received a website with a flat menu structure. Metacognitive support through a checklist did not enhance learning outcomes. H4 has to be rejected. Furthermore, a checklist did not increase ICL or GR, even if ECL was reduced, H5 could not be supported. So far, the results suggested (or at least gave hints) that the deep menu structure might be suppressive for certain learning processes. Nevertheless, including a checklist in the conditions with a deep menu structure enhanced retention performance (but not transfer performance). The checklist could further compensate for the negative effect of the deep menu structure regarding perceived ECL, but no interaction effects were found for productive load. Consequently, H6 could only be supported partially. In contrast to H7, including a checklist did not influence metacognitive judgments as well as accuracy scores in general. Furthermore, there were no interaction effects regarding the experimental factors. Consequently, H8 has to be rejected as well. The explorative research questions offered further interesting insights into learning processes. Even if learning outcomes did not differ with regard to the menu structure, learners receiving the deep menu structure learned significantly longer than learners receiving a flat menu structure. Consequently, instructional efficiency decreased in the deep menu structure condition. However, including a checklist could compensate for this negative effect at least regarding the efficiency score for the knowledge acquisition, but not regarding the efficiency score derived from transfer items. Furthermore, the checklist was perceived as particularly effective in the condition with a deep menu structure, further outlining how a checklist might compensate negative effects of a suppressing menu structure.

At first, the effects regarding menu structure have to be discussed. Learning outcomes were not influenced through the experimental manipulation. Consequently, the beneficial effects of additional segmentation might be at least compensated through the need for additional navigation to get to the information of the course. This negative effect of navigational needs is reflected in two important measures. First, in line with CLT (Sweller et al., 2019), ECL was increased in the conditions with a deep menu structure. Navigating through web pages thus induced unproductive load, which suppressed learning. Consequently, the beneficial effect of segmenting could not unfold. Furthermore, the need for navigation was reflected in an increase in unproductive learning time. The additional learning time was invested in navigating through the web pages (to deal with the unproductive load) and not in schema construction processes. Nevertheless, including a checklist compensated negative effects with regard to unproductive load as well as instructional efficiency. Therefore, it is important to discuss that, in contrast to the theoretical assumptions, these compensating effects were not based or reflected in metacognitive processes. The compensating effects might rather be explained considering cognitive processes or metacognitive factors that were not assessed through our measures. A checklist might be used for cognitive offloading and might lead to a change in the general learning strategy. Learners had a constantly accessible possibility to review their learning progress. Consequently, web pages were called up in a more targeted manner and information did not have to be searched again because of uncertainty as to whether it was fully received. As a result, unproductive learning time as well as perceived unproductive load was lowered and the checklist was perceived as particularly facilitating when information has to be searched due to a deep menu structure. This change in strategy might be adapted in an early stage of the learning process and rather influenced concrete cognitive learning processes rather than metacognitive judgments. Modifying the checklist by not asking whether information was already read or learned but by asking how well the sub-topics were understood (comparable with JOL ratings) might actually trigger metacognitive processes measurable by the used measures. Nevertheless, including a checklist might compensate negative effects of a suppressing menu structure, but was not beneficial for learning in general. In line with the basic assumptions of the CLT (Sweller et al., 2019), learners do not need instructional support when cognitive capacity is sufficient for processing the learning content. In the conditions with a flat menu structure, learners might not be heavily loaded as no navigational needs were induced. Thus, no change in learning strategy was necessary as enough cognitive resources were available for processing the learning information as well as dealing with the website. Instructional support only became important when the structure of the website or the structure of the information problems exceeded cognitive capacity.

4.1. Implications

On theoretical side, it became clear how much different research disciplines can benefit from each other. The segmenting principle as an instructional design principle might serve as an explanation for learning effects with different menu structures, but it becomes clear that manipulating a website menu changed more than just the segmentation of information. Consequently, research with regard to website usability could enhance the understanding of learning with different menu structures. In particular, with regard to multimedia learning, considering different disciplines led to a more differentiated picture of learning processes. Furthermore, the results showed how even simple support options like checklists influence learning strategies and compensate negative effects of suppressing multimedia learning designs. Researchers should be encouraged to investigate how, under which conditions instructional support facilitates learning and, in particular, how to measure the change in the concrete learning processes.

On practical side, designers should be aware that the menu structure affects learning processes. At first sight, it might be intuitively beneficial for learners to place every information at a separate spot within the website menu, to ensure a clear structure and order, but usability and instructional efficiency can suffer greatly from such an approach. Nevertheless, the instructional information was rather simple in the current study, and when instructing more complex information, instructors may have no choice but to create a deeper menu structure. Consequently, designers of e-learning websites should keep the balance between an information presentation that does not overload learners due to the amount of presented information and a menu that does not overburden them because of the need for navigation. Instructional support through a checklist might be an additional tool which is worth considering. Designing a website with complex instructional information should benefit from the support that leads to cognitive offloading or that induces beneficial learning strategies. Therefore, a checklist is an easy-to-implement tool to achieve these goals.

4.2. Limitations

The inclusion of a checklist could compensate for learning inhibiting effects of a deep menu structure, but metacognitive measures failed to show beneficial effects. As the checklist might indeed have effects on metacognitive strategies since the early stage of learning, additional measures like the ease of learning judgments (Son and Kornell, 2008) would have provided further insights. Furthermore, the current study investigated menu structure by presenting all information at once vs. providing one additional menu level. E-learning websites can have further menu levels, which were not considered here. Finally, the study was carried out with a (mostly female) student sample. Usually, students have a broad experience navigating through website menus as most websites (e.g., information databases, shopping websites) require a menu-based navigation. Even if it is particularly interesting that learning-relevant effects could still be found, generalizability is limited.

5. Conclusions

Overall, the study has achieved the objectives as follows: It could be outlined that a deep menu structure is unfavorable for e-learning websites. It is possible to add additional instructional support to compensate for this negative effect, but the menu structure should be considered fundamentally to increase the instructional quality. Future research should consider measuring additional metacognitive variables; for example, ease of learning judgments that further specify how learning is perceived at the beginning of the process and monitored further on. It is additionally possible that results change (or further intensify) if additional menu levels are added. For example, the missing effect of productive load might be explained by the fact that the need to navigate only through one level might be sufficient to influence unproductive load, but suppressing effects on productive load only become visible in more complex menu structures. Further research might manipulate the menu structure by adding additional levels. Manipulating menu levels in consistently ascending order (one level vs. two levels, vs. three levels, vs. …) even allows a regression-analytical evaluation of the results. With respect to the acquired (mostly) female sample, other samples should be considered in future research. For example, secondary (or even primary) students might be less experienced in working with menu-based websites or less experienced in working with digital devices in general. Considering this sample, the need for additional navigation might be even more harmful and the role of instructional support might be even more important.

Data availability statement

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

Ethics statement

Ethical review and approval was not required for the study on human participants in accordance with the local legislation and institutional requirements. The patients/participants provided their written informed consent to participate in this study.

Author contributions

MB: conceptualization, methodology, investigation, data curation, formal analysis, validation, project administration, writing—original draft, writing—review and editing, and supervision. DS: formal analysis, writing—original draft, and writing—review and editing. EW: conceptualization, methodology, investigation, data curation, writing—review, and editing. All authors contributed to the article and approved the submitted version.

Funding

The article processing charge was funded by the Baden-Württemberg Ministry of Science, Research and Culture and the University of Education, Freiburg in the funding program Open Access Publishing.

Acknowledgments

We want to thank Anastasia Markov, Adriana Richter, and Patricia Berthold for their contribution to this study. Under the supervision of the corresponding author, they worked out the prototypical study idea, designed the study materials, and conducted the first part of the survey as part of a university course on scientific work.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

Akturk, A. O., and Sahin, I. (2011). Literature review on metacognition and its measurement. *Proc. Soc. Behav. Sci.* 15, 3731–3736. doi: 10.1016/j.sbspro.2011.04.364

Beege, M., Krieglstein, F., Schneider, S., Nebel, S., and Rey, G. D. (2021). Is there a (dis-) fluency effect in learning with handwritten instructional texts. Evidence from three studies. *Front. Educ.* 6, 678798. doi: 10.3389/feduc.2021.678798

BPS Bildungsportal Sachsen GmbH (2022). *OPAL*. Available online at: https://bildungsportal.sachsen.de/opal/home?1 (accessed January 13, 2023).

Cockburn, A., Gutwin, C., and Greenberg, S. (2007). "A predictive model of menu performance," in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (San Jose, CA), 627–636.

Cook, D. A., and Dupras, D. M. (2004). A practical guide to developing effective web-based learning. *J. Gen. Intern. Med.* 19, 698–707. doi: 10.1111/j.1525-1497.2004.30029.x

Corpuz-Abenoja, M. M. R. (2022). Task delegation and completion checklist (TDCC) in science research team project. *Eur. J. Educ. Stud.* 9, 311–328. doi: 10.46827/ejes.v9i2.4185

Dinsmore, D. L., and Parkinson, M. M. (2013). What are confidence judgments made of? Students' explanations for their confidence ratings and what that means for calibration. *Learn. Instruct.* 24, 4–14. doi: 10.1016/j.learninstruc.2012.06.001

Douglas, D. E., and Van Der Vyver, G. (2004). Effectiveness of e-learning course materials for learning database management systems: an experimental investigation. *J. Comp. Inf. Syst.* 44, 41–48.

Dunlosky, J., and Thiede, K. W. (2013). "Metamemory," in *The Oxford Handbook of Psychology*, ed D. Reisberg (Oxford: Oxford University Press), 283–298.

Erdfelder, E., Faul, F., and Buchner, A. (1996). GPOWER: a general power analysis program. *Behav. Res. Methods Instrum. Comp.* 28, 1–11. doi: 10.3758/BF03203630

Farris, J. S., Jones, K. S., and Elgin, P. D. (2002). Users' schemata of hypermedia: what is so "spatial" about a website? *Interact. Comput.* 14, 487–502. doi: 10.1016/S0953-5438(02)00011-5

Flavell, J. H. (1979). Metacognition and cognitive monitoring: a new area of cognitive–developmental inquiry. *Am. Psychol.* 34, 906–911. doi: 10.1037/0003-066X.34.10.906

Geven, A., Sefelin, R., and Tscheligi, M. (2006). "Depth and breadth away from the desktop: the optimal information hierarchy for mobile use," in *Proceedings of the 8th Conference on Human-Computer Interaction With Mobile Devices and Services* (Helsinki), 157–164.

Guo, L. (2022). Using metacognitive prompts to enhance self-regulated learning and learning outcomes: a meta-analysis of experimental studies in computer-based learning environments. *J. Comp. Assist. Learn.* 38, 811–832. doi: 10.1111/jcal.12650

Hasler, B. S., Kersten, B., and Sweller, J. (2007). Learner control, cognitive load and instructional animation. *Appl. Cogn. Psychol.* 21, 713–729. doi: 10.1002/acp.1345

Kalyuga, S., and Singh, A. M. (2016). Rethinking the boundaries of cognitive load theory in complex learning. *Educ. Psychol. Rev.* 28, 831–852. doi: 10.1007/s10648-015-9352-0

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

Koo, T. K., and Li, M. Y. (2016). A guideline of selecting and reporting intraclass correlation coefficients for reliability research. *J. Chiropr. Med.* 15, 155–163. doi: 10.1016/j.jcm.2016.02.012

Krieglstein, F., Beege, M., Rey, G. D., Ginns, P., Krell, M., and Schneider, S. (2022). A systematic meta-analysis of the reliability and validity of subjective cognitive load questionnaires in experimental multimedia learning research. *Educ. Psychol. Rev.* 34, 2485–2541. doi: 10.1007/s10648-022-09683-4

Kurtenbach, G. P. (1993). *The Design and Evaluation of Marking Menus*, 1–173. Toronto, ON: University of Toronto.

Lee, M. G. (2001). Profiling students' adaptation styles in Web-based learning. *Comp. Educ.* 36, 121–132. doi: 10.1016/S0360-1315(00)00046-4

Mayer, R. E. (2014). "Introduction to multimedia learning," in *The Cambridge Handbook of Multimedia Learning, 2nd Edn.*, ed R. E. Mayer (Cambridge: Cambridge University Press), 1–24.

Mayer, R. E. (2021). "Cognitive theory of multimedia learning," in *The Cambridge Handbook of Multimedia Learning (Cambridge Handbooks in Psychology)*, eds R. Mayer, and L. Fiorella (Cambridge: Cambridge University Press), 57–72.

Mayer, R. E., and Pilegard, C. (2014). "Principles for managing essential processing in multimedia learning: segmenting, pre-training, and modality principles," in *The Cambridge Handbook of Multimedia Learning, 2nd Edn*, ed R. E. Mayer (Cambridge: Cambridge University Press), 316–344.

McGuinness, C. (1990). "Talking about thinking: the role of metacognition in teaching thinking," in *Lines of Thinking*, eds K. Gilhooly, M. Deane, and G. Erdos (San Diego, CA: Academic), 310–312.

Moreno, R., and Mayer, R. E. (2007). Interactive multimodal learning environments. *Educ. Psychol. Rev.* 19, 309–326. doi: 10.1007/s10648-007-9047-2

Mudrick, N. V., Azevedo, R., and Taub, M. (2019). Integrating metacognitive judgments and eye movements using sequential pattern mining to understand processes underlying multimedia learning. *Comput. Human Behav.* 96, 223–234. doi: 10.1016/j.chb.2018.06.028

Paas, F., and van Merriënboer, J. J. G. (2020). Cognitive-load theory: methods to manage working memory load in the learning of complex tasks. *Curr. Dir. Psychol. Sci.* 29, 394–398. doi: 10.1177/0963721420922183

Patsula, P. J., Detenber, B. H., and Theng, Y. L. (2010). Structure processing of web-based menus. *Int. J. Hum. Comp. Interact.* 26, 675–702. doi: 10.1080/10447318.2010.487197

Pham, M., Singh, K., and Jahnke, I. (2021). Socio-technical-pedagogical usability of online courses for older adult learners. *Interact. Learn. Environ.* 1–17. doi: 10.1080/10494820.2021.1912784

Pieger, E., Mengelkamp, C., and Bannert, M. (2016). Metacognitive judgments and disfluency–Does disfluency lead to more accurate judgments, better control, and better performance? *Learn. Instruct.* 44, 31–40. doi: 10.1016/j.learninstruc.2016.01.012

Prezenski, S., and Russwinkel, N. (2014). Combining cognitive ACT-R models with usability testing reveals users mental model while shopping with a smartphone application. *Int. J. Adv. Intell. Syst.* 7, 700–715.

Rey, G. D., Beege, M., Nebel, S., Wirzberger, M., Schmitt, T. H., and Schneider, S. (2019). A meta-analysis of the segmenting effect. *Educ. Psychol. Rev.* 31, 389–419. doi: 10.1007/s10648-018-9456-4

Rowlands, K. D. (2007). Check it out! Using checklists to support student learning. *Eng. J.* 96, 61–66.

Samp, K. (2013). "Designing graphical menus for novices and experts: connecting design characteristics with design goals," in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Paris), 3159–3168.

Schnotz, W., and Lowe, R. (2008). "A unified view of learning from animated and static graphics," in *Learning With Animations: Research Implications for Design*, eds R. Lowe, and W. Schnotz (New York, NY: Cambridge University Press), 304–356.

Schraw, G. (2009). A conceptual analysis of five measures of metacognitive monitoring. *Metacogn. Learn.* 4, 33–45. doi: 10.1007/s11409-008-9031-3

Snowberry, K., Parkinson, S. R., and Sisson, N. (1983). Computer display menus. *Ergonomics* 26, 699–712. doi: 10.1080/00140138308963390

Son, L. K., and Kornell, N. (2008). "Research on the allocation of study time: key studies from 1890 to the present (and beyond)," in *Handbook of Metamemory and Memory*, eds J. Dunlosky, and R. A. Bjork (New York, NY, US: Psychology Press), 333–351.

Son, L. K., and Metcalfe, J. (2005). Judgments of learning: evidence for a two-stage process. *Mem. Cognit.* 33, 1116–1129. doi: 10.3758/BF03193217

Spanjers, I. A. E., Van Gog, T., and Van Merrienboer, J. J. G. (2010). A theoretical analysis of how segmentation of dynamic visualizations optimizes students' learning. *Educ. Psychol. Rev.* 22, 411–423. doi: 10.1007/s10648-010-9135-6

Stiller, K. D., Petzold, K., and Zinnbauer, P. (2011). Presentation time concerning system-paced multimedia instructions and the superiority of learner pacing. *Aust. J. Educ. Technol.* 27, 693–708. doi: 10.14742/ajet.945

Sweller, J. (2020). Cognitive load theory and educational technology. *Educ. Technol. Res. Dev.* 68, 1–16. doi: 10.1007/s11423-019-09701-3

Sweller, J., van Merriënboer, J. J., and Paas, F. (2019). Cognitive architecture and instructional design: 20 years later. *Educ. Psychol. Rev.* 31, 261–292. doi: 10.1007/s10648-019-09465-5

Tabbers, H. K., and de Koeijer, B. (2010). Learner control in animated multimedia instructions. *Instruct. Sci.* 38, 441–453. doi: 10.1007/s11251-009-9119-4

Thiede, K. W., Anderson, M. C. M., and Therriault, D. (2003). Accuracy of metacognitive monitoring affects learning of texts. *J. Educ. Psychol.* 95, 66–73. doi: 10.1037/0022-0663.95.1.66

Toan, P. N., Dang, T. T., and Hong, L. T. T. (2021). E-learning platform assessment and selection using two-stage multi-criteria decision-making approach with grey theory: a case study in Vietnam. *Mathematics* 9, 3136. doi: 10.3390/math9233136

Ukrayinska, O. (2020). "Using checklists for developing student teachers' language assessment literacy," in *Perspectives on Language Assessment Literacy*, ed S. Hidri (London: Routledge), 84–106.

Van Gog, T., and Paas, F. (2008). Instructional efficiency: Revisiting the original construct in educational research. *Educ. Psychol.* 43, 16–26. doi: 10.1080/00461520701756248

Wickens, C. D., Hollands, J. G., Banbury, S., and Parasuraman, R. (2013). *Engineering Psychology and Human Performance, 4th Edn.* Boston, MA: Pearson Education.

Wouters, P. (2007). *How to Optimize Cognitive Load for Learning From Animated Models* (Doctor of Philosophy). Open University of the Netherlands, Maastricht, Netherlands.

Zheng, L. (2016). The effectiveness of self-regulated learning scaffolds on academic performance in computer-based learning environments: a meta-analysis. *Asia Pac. Educ. Rev.* 17, 187–202. doi: 10.1007/s12564-016-9426-9