Learning and Instruction 94 (2024) 101988

Contents lists available at ScienceDirect

Learning and Instruction

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

Does watching an explainer video help learning with subsequent text? – Only when prompt-questions are provided

Marie-Christin Krebs *, Katharina Braschoß, Alexander Eitel

Department of Educational Psychology, Justus-Liebig-Universität Giessen, Otto-Behaghel-Str. 10, Giessen, 35390, Germany

A R T I C L E I N F O

Keywords:
Explainer video
Cognitive prompts
Video-based learning
Illusion of understanding
Situational interest

A B S T R A C T

Background: Learning with explainer videos can foster learning. However, their effects on subsequent learning are still unclear. On the one hand, they might increase situational interest and scaffold subsequent learning. On the other hand, they might hinder subsequent learning by fostering an illusion of understanding. In case of the latter, the question arises of whether providing prompt-questions after an explainer video would prevent an illusion of understanding. Therefore, we investigated the effects of medium and prompt-questions on subsequent learning with text.
Sample: One hundred thirty-three teacher students and psychology students from a German university.
Methods: In an online study with a 2x2 between-subjects design, we investigated the effects of medium (video vs. video-script) in learning phase 1 and prompt-questions (yes vs. no) on subsequent learning with text.
Results: As expected, watching the video made the content seem more interesting and less difficult. Contrary to the illusion-of-understanding-assumption, this did not result in learners overestimating but rather underestimating themselves. Moreover, while prompt-questions in the video condition fostered learning, they impaired learning in the video-script condition. Exploratory mediation analyses revealed that in the prompt condition, the superiority of the video was mainly driven by the quality of the prompt-answers rather than the time learners invested in answering the prompt-questions.
Conclusions: Our findings suggest that explainer videos combined with prompt-questions can foster learning with subsequent text. However, further research is necessary to replicate the findings under more controlled conditions and to investigate the underlying processes in greater depth.

Learning with videos is popular. Millions of learners watch so-called explainer videos (i.e., short instructional videos, which explain complex/abstract issues in an entertaining way) on all kinds of topics. It is therefore not surprising that there has been growing interest recently in whether explainer videos effectively support learning. While there is evidence that videos can be an effective instructional tool, they are not always learning-effective, a fact we have known for almost 40 years (e.g., Salomon, 1984).

To make learning with videos effective, the video needs to be 1) well designed and 2) well implemented within an instruction. The first aspect has been the focus of ample previous research, resulting in various design principles (e.g., Brame, 2016; Kulgemeyer, 2020; Mayer, 2021). However, also highly relevant are the questions of how videos should be implemented, for example in learning environments also containing other types of learning materials (e.g., texts), and what problems might emerge in this context. For instance, learners preparing for an exam should not just rely on a short explainer video but also review other materials such as their notes or a textbook chapter. However, relatively little is currently known about whether and how watching explainer videos interacts with subsequently learning from text.

Based on previous research, there are – among others – two possibilities how watching an explainer video first might influence subsequent learning with text. On the one hand, watching the explainer video first might foster subsequent learning from text by raising situational interest (Endres et al., 2020) and/or by providing a mental scaffold (Eitel & Scheiter, 2015). On the other hand, watching the explainer video first might hinder learning from text by fostering an illusion of understanding (Kulgemeyer & Wittwer, 2023; Salomon, 1984), where learners prematurely think that they have understood the topic well enough (Kulgemeyer & Wittwer, 2023; Wittwer & Renkl, 2008). This

* Corresponding author.
E-mail addresses: marie-christin.krebs@psychol.uni-giessen.de (M.-C. Krebs), katharina.braschoss@psychol.uni-giessen.de (K. Braschoß), alexander.eitel@psychol.uni-giessen.de (A. Eitel).

https://doi.org/10.1016/j.learninstruc.2024.101988
Received 31 March 2023; Received in revised form 21 July 2024; Accepted 25 July 2024
Available online 4 August 2024

would reduce time and effort to process subsequent text. In case of the latter, the question arises of how to prevent such an illusion of understanding. One means may be cognitive and metacognitive prompts, that is, content-related and monitoring-related questions (e.g., Berthold et al., 2007) provided after watching the video. There is evidence that such prompts benefit learning from cognitive and metacognitive perspectives because they serve as a retrieval practice task (e.g., Roediger & Karpicke, 2006), and foster metacognitive monitoring (e.g., Fiorella et al., 2020; Müller & Seufert, 2018; Szpunar et al., 2014). Accordingly, they could also foster learning with explainer videos with subsequent expository text by stimulating beneficial learning strategies.

Against this backdrop, the aim of the present study was to shed light on the question of whether explainer videos hamper subsequent learning by inducing an illusion of understanding, and whether providing learners with prompts might counter possible negative effects of explainer videos on subsequent learning.

1. Learning with explainer videos – is it as simple as it seems?

Explainer videos, also referred to as explanatory videos (e.g., Kulgemeyer, 2020) or educational videos (Brame, 2016), are short audio-visual presentations (e.g., animated videos, screencasts, sketched explanation videos) that cover all kinds of educational topics and can thus be considered a subcategory of instructional videos (i.e. audio-visual presentations intended to promote learning; Mayer, 2021). Explainer videos explain relatively complex and abstract issues in relatively short time, conversational language, and usually in an entertaining way (e.g., Krämer & Böhrs, 2017). Explainer videos can be roughly divided into two categories: explainer videos that present more abstract concepts and conceptual (declarative) knowledge, and explainer videos that convey how-to (procedural) knowledge. The latter category is usually referred to as video tutorials and is considered a subcategory of explainer videos (Wolf, 2015). Importantly, explainer videos do not exhaustively explain a topic, but rather highlight relevant issues. In educational contexts, learners watch explainer videos, for example, to review learning content they did not understand in class or to prepare for exams (Wolf et al., 2021). But do explainer videos support learning generally, or do learners merely succumb to this illusion?

On the one hand, there is empirical evidence that video tutorials are more effective for learning than text tutorials (e.g., Lloyd & Robertson, 2012; van der Meij & van der Meij, 2014). For instance, van der Meij and van der Meij (2014) found positive effects of tutorials for software learning. They compared four types of tutorials that included either paper-based or video-based instruction and were accompanied by either a short paper-based or video-based preview, resulting in the following conditions: Paper-based preview and paper-based instruction, paper-based preview and video procedure, video preview and paper-based procedure, video preview and video procedure. The paper-based or video-based preview defined the problem and displayed the start and end screen of the task. With regard to post-test performance, results indicated that learners with a video preview outperformed learners in the plain-text conditions.

On the other hand, there is also empirical evidence that for declarative knowledge expository texts are more effective as a first information source compared to videos. List and Ballenger (2019) investigated the effect of two complementary media (text vs. video) on strategy use and comprehension. All learners received a more expository source first (text vs. video) before they received a second more narrative source (text vs. video). Results indicated that learners differed regarding their strategy use based on the medium they received. For instance, learners in the text only condition reported more cross-textual elaboration as well as organizational strategies. Moreover, learners who received a text first scored significantly higher on a knowledge post-test.

Results of a recent meta-analysis by Noetel et al. (2021) showed that providing learners with supplemental videos in addition to existing learning content facilitated student learning ($g > 0.2$). However, their meta-analysis summarised effects from all kinds of videos (e.g., lecture capture, educational multimedia, tutorials), and not just explainer videos. Other research has indicated positive effects of animated *explainer* videos on self-assessed outcome variables such as interest, engagement, and understanding a topic, but not on objective learning outcomes such as practical and written exam scores (Liu & Elms, 2019). Hence, the evidence is mixed, potentially because the video design and implementation must be considered.

According to Mayer (2021), learning with instructional videos is effective for learning if they are designed in line with three principles, namely the dual channels principle (Paivio, 1986), limited capacity principle (Sweller et al., 2011), and generative activity principle (Mayer & Fiorella, 2022). Brame (2016) considered cognitive load, learner engagement, and active learning as the three most important factors for designing and implementing explainer videos effectively. Kulgemeyer (2020) postulated a framework for science explainer videos, which also included several design recommendations (e.g., using minimal explanation, follow-up learning tasks, focussing on complex scientific principles, etc.). Kulgemeyer (2020) tested this framework by comparing a science explainer video developed with a high vs. a low fit to the postulated framework. It was found that the learners with an explainer video fitting the framework outperformed those with the less-fitting explainer video regarding their declarative knowledge, thus further supporting the assumption that learners can benefit from well-designed explainer videos. What "well-designed" and "well-implemented" means in particular, requires an answer from *three* different perspectives that are not mutually exclusive, but probably complement each other.

From a *motivational* perspective, explainer videos may foster learning especially when they are visually appealing and comprise a personalised frame story (e.g., Endres et al., 2020). They may thus make the learning experience more joyful due to their "emotional" design, which sometimes leads to better learning outcomes (Brom et al., 2018). Specifically, explainer videos comprising emotional design elements lead to better-sustained attention during learning when they trigger and maintain learners' situational interest (e.g., Endres et al., 2020; Hidi & Renninger, 2006). In contrast to individual interest, which is defined as a relatively stable personal characteristic, situational interest is described as a reaction to environmental characteristics (Hidi, 2001). Hence, situational interest can be influenced by the design characteristics of learning materials (Endres et al., 2020). While triggered situational interest refers to superficial features of the instruction that might shortly catch learners' attention (Endres et al., 2020; Hidi, 1990), maintained situational interest is responsible for maintaining focused attention and persistence over an extended learning phase (Hidi & Renninger, 2006). According to Endres et al. (2020), maintained situational interest should yield desirable effects on learning, particularly when learners are required to engage with the learning material in a longer learning phase.

From a *cognitive* perspective, explainer videos may help learners to gain initial conceptual and/or procedural understanding, which serves as a kind of *cognitive scaffold* for further learning (e.g., Eitel & Scheiter, 2015). In the context of multimedia learning, for example, results of a review by Eitel and Scheiter (2015) suggest that learners benefit from receiving the medium with the less complex information first. According to Eitel and Scheiter (2015), this could be because processing the less complex information in the first medium might facilitate the processing of the more complex information in the second medium. Consequently, watching an explainer video that briefly summarises relevant information of a topic in a structured overview might also facilitate subsequent learning processes with other media containing more complex information (e.g., a text). Moreover, explainer videos may be even more effective as a cognitive scaffold compared to text.

According to the Cognitive Theory of Multimedia Learning (CTML; Mayer, 2022), textual and visual information is processed in two different channels. Therefore, watching an explainer video that contains visual as well as auditory elements can result in a richer mental representation than solely reading a text with the same information (i.e. multimedia effect; Mayer, 2022). On the other hand, there is also empirical evidence that text might be the more facilitative medium for information presentation, and thus should be presented first (List & Ballenger, 2019). Overall, it is therefore an open question whether learners would benefit more from watching an explainer video compared to reading the same information as a text before learning with an expository text.

From a *meta-cognitive perspective*, breaking down complex information and presenting them in an entertaining way via video might lead to an *illusion of understanding* (Kulgemeyer & Wittwer, 2023; List & Ballenger, 2019; Salomon, 1984), meaning that watching the (explainer) video is easy, so understanding the contents may appear easier than it actually is. As a consequence, learners may overestimate their level of understanding (Kulgemeyer & Wittwer, 2023; Wittwer & Renkl, 2008), and thus invest (too) little mental effort in further processing the learning content, which could negatively affect their learning outcomes (e.g., Paik & Schraw, 2013). Moreover, Kornell et al. (2011) found that the perceived ease of information processing could bias learners' metacognitive judgements (i.e. ease-of-processing heuristic; Kornell et al., 2011). If explainer videos are easy to process or at least appear to be so, this might result in learners overestimating their actual knowledge level (i.e. metacognitive calibration), because they base their judgement of expected performance on the perceived easiness of processing/memorizing the video content (i.e. data-driven self-regulation; Baars et al., 2020). Additionally, learners tend to overestimate their learning when learning with multimedia material (i.e. multimedia heuristic; Eitel, 2016; Hoch et al., 2023; Serra & Dunlosky, 2010). Since videos can be considered multimedia learning material, the multimedia heuristic might also apply to learning with videos. It is possible that learners overestimate their knowledge level after watching the explainer video because they expect that learning with explainer videos is easy.

In the case of explainer videos, learners may also find subsequent learning material (e.g., book chapters) irrelevant and/or unnecessarily complicated compared to the explainer video, and thus fail to engage in further learning processes. Previous research suggests that this is potentially especially problematic if explainer videos contain errors or convey scientific misconceptions, because misconceptions are close to everyday experiences, and thus potentially more attractive than the scientifically correct explanations (Kulgemeyer & Wittwer, 2023). Moreover, Senko et al. (2022) found that situational interest promotes learners' overconfidence. This too could be especially problematic for explainer videos, as they are usually designed to promote situational interest. Yet this is precisely what can lead to learners being less capable of correctly assessing their actual knowledge level and of adapting their further learning process accordingly. Moreover, as watching a video is a rather passive activity, learners may not engage cognitively with the content presented to them. Consequently, they might not even become aware of their illusion of understanding (Kulgemeyer & Wittwer, 2023).

Metacognitive monitoring is a key factor for guiding learners' control of subsequent learning behaviour (e.g., Dunlosky & Rawson, 2012). For instance, there is empirical evidence that accurate metacognitive judgements improve learning performance (e.g., Ackerman & Leiser, 2014; Dunlosky & Rawson, 2012). However, explainer videos might make it harder for learners to monitor their knowledge level accurately. Consequently, this might have negative effects on subsequent learning processes.

To the best of our knowledge, up to date there is only little empirical research on the differences regarding metacognitive calibration between text and video as single learning materials, let alone on possible effects on subsequent learning phases with expository texts (e.g., Mason et al., 2022; Tarchi et al., 2021; Tarchi & Mason, 2022). Tarchi et al. (2021) compared learning from text with learning from instructional videos and subtitled instructional videos. In this study, results did not indicate a significant difference between conditions regarding metacognitive calibration, although learners reported a higher perceived self-efficacy when learning from video compared to text. Moreover, results did not indicate an effect of medium on immediate post-test performance but an advantage of the text condition compared to the subtitled video condition in a delayed post-test. In another study with second language learners, Tarchi and Mason (2022) also compared the effect of text, instructional video and subtitled instructional video on learning. Here, learners in the video condition outperformed learners in a text condition in a delayed post-test, but again there were no significant differences between the different media in an immediate post-test. Moreover, there were also no significant differences regarding the self-rated judgements on their performance between the media.

In a recent study, Mason et al. (2022) also compared the effect of medium and context on learning with multiple texts versus multiple instructional videos. Results showed that learners in the video condition invested longer learning time compared to those in the text condition. However, there were no significant effects of medium or context on learners' post-test performance. Overall, empirical evidence on the effect of learning with multiple media (i.e. explainer video *and* text) on learning processes and learning outcomes is still scarce and findings from single medium studies are mixed.

However, if explainer videos might be especially prone to foster inaccurate metacognitive judgements, how can we encourage learners to engage more actively with the content and more accurately assess their current level of understanding to counter potential negative effects on subsequent learning?

1.1. Using prompts to counter the illusion of understanding

Previous research has shown that prompts can assume multiple forms (e.g., hints, questions) and, as a subcategory of scaffolds, can support both self-regulated learning and learning performance (for an overview, see Zheng, 2016).

According to Berthold et al. (2007), prompts can serve as strategy activators, i.e., they can elicit beneficial learning strategies that learners are capable of but do not spontaneously and/or adequately demonstrate. For instance, Berthold et al. (2007) used writing instructions with cognitive prompts, metacognitive prompts, or a mixture of both to elicit cognitive and metacognitive strategies in writing learning protocols. Their results showed that the mixture of cognitive and metacognitive prompts had positive effects on learning outcomes. Further, providing learners with only cognitive prompts not only elicited cognitive learning strategies, it also triggered significantly more metacognitive strategies. Prompts are thus effective both as strategy activators and as monitoring aids because they stimulate the comparison of one's current learning state with one's actual learning goal, and thus reveal knowledge gaps or illusions of understanding (Müller & Seufert, 2018).

Furthermore, content-related questions after watching an explainer video might serve as *retrieval practice task* (for a review, see Agarwal et al., 2021), and thus help learners to consolidate their knowledge - an important part of the learning process (Roelle et al., 2022). Szpunar et al. (2013), for example, demonstrated that administering tests after each video segment of an online video-lecture significantly reduced students' mind wandering, and resulted in better learning outcomes compared to giving a test only after the last video segment. In a similar study, Szpunar et al. (2014) showed that students again benefitted from repeated testing during a video-lecture: They tended to be significantly less likely to underestimate or overestimate their knowledge and performed better in a post-test. Overall, these findings suggest that content-related questions can promote metacognitive monitoring and learning outcomes.

Furthermore, in line with the generative activity principle (Mayer & Fiorella, 2022) content-related questions might also serve as *generative learning task*. Generative learning activities (e.g., self-explanations, drawings, etc.) encourage learners to cognitively (re)engage with the learning material, fostering deeper processing of the content, and thus the construction of coherent mental representations, resulting in better comprehension and transfer (Fiorella et al., 2020; Mayer & Fiorella, 2022; Roelle et al., 2022). Accordingly, Wang et al. (2023) showed that prompting learners to engage in a generative learning activity (i.e. writing a summary) between video segments benefitted learning.

According to Roelle et al. (2022), both retrieval practice tasks and generative learning tasks are beneficial for learning, because both tasks can support self-monitoring, and thus affect subsequent learning. Against this background, content-related and monitoring-related questions after watching the explainer video may 1) support learners in constructing and consolidating coherent mental representations of the new knowledge, 2) provide implicit feedback about the learners' current level of understanding, and thus 3) support learners in adapting their later learning processes accordingly. Therefore, we expected that content-related and monitoring-related questions would prevent learners from developing an illusion of understanding after watching the explainer video. Moreover, we expected that all learners in our study would benefit from these prompts in terms of their learning performance and outcome.

1.2. The present study and hypotheses

We investigate if prompts might support learning from explainer videos before text. To do so, we compared learning processes and outcomes of students who learned with explainer videos or the corresponding video scripts, and received prompts or not, before they should learn with a textbook. We formulated and preregistered (AsPredicted#84960) the following (competing) hypotheses.

H1. Situational Interest Hypothesis

Explainer videos may trigger situational interest. Therefore, learners in the video condition should report higher triggered situational interest in the first learning phase compared to the video-script condition (H1a). Moreover, the higher triggered situational interest for learners in the video condition in the first learning phase should result in higher maintained situational interest in the second learning phase and accordingly in learners investing more learning time in the second learning phase, and thus performing better at the post-test (H1b).

H2. Illusion of Understanding Hypothesis

Explainer videos without prompts may trigger an illusion of understanding which might be reflected in more passive processing behaviour. Accordingly, learners watching the explainer video should report lower judgements of active as well as passive mental effort and lower judgements of difficulty after the first learning phase compared to learners in the video-script condition (H2a). Moreover, we expected that learners in the video condition without prompts would report higher judgements of learning in the first learning phase compared to actual learning outcomes in the post-test (i.e. metacognitive calibration) compared to the video-script condition without prompts (H2b). Additionally, we expected that an illusion of understanding in phase 1 would lead learners in the video condition without prompts to invest less learning time in the second learning phase, which in turn would result in poorer learning outcomes (H2c).

H3. Prompts Support Learning Process Hypothesis

If prompts foster learning with explainer videos because they help learners assess their knowledge more accurately, learners with prompts (vs. no prompts) should have more accurate metacognitive judgements (H3a) and invest higher mental effort (H3b) and more learning time (H3c) in the second learning phase.

H4. Prompts Support Learning Outcomes Hypothesis

We expected that all learners, but especially those with explainer videos, would benefit from prompts. Thus, learners with prompts should outperform those without in the post-test (H4a). In addition, learners in the video-prompt condition should perform best (H4b).

1.3. Exploratory analyses

Against the background that explainer videos may both arouse situational interest and serve as a cognitive scaffold, we also investigated whether learners who watched the video (vs. read the video-script) would be better at answering the prompt questions after the first learning phase, and whether higher-quality answers would mediate the prompts' effect on learning outcomes.

2. Methods

2.1. Participants and design

Initially, 141 teacher students and psychology students from a German university participated in our online learning experiment. Four participants reported not having faithfully followed the instruction during the study and were thus excluded from data analyses. Another person was excluded as they had a lot of prior knowledge about the learning topic, and another three because they reported having been distracted while learning and also having used external resources for the post-test. After exclusion, 133 participants (108 female; 25 male; $M = 21.54$ years, $SD = 4.42$) remained. We conducted an a priori power analysis using G*Power 3.1 (Faul et al., 2009) for an assumed medium effect size (alpha = 0.05; power = 0.80, $f = 0.25$), which resulted in a recommended sample size of $N = 128$ participants.

The design of the online experiment was a 2x2 between-subjects design with medium in the first learning phase (video vs. video-script) and prompts after the first learning phase (no prompts vs. prompts) as independent variables. Participants were randomly assigned to one of four conditions: video-script without prompts (n = 35), video-script with prompts (n = 22), video without prompts (n = 46), and video with prompts (n = 30). The median time to complete the study was 60 min. Participants could win one of 30 gift cards (10€) as compensation and were recruited during lectures and courses for teacher students. The study was approved by our local ethics board (LEK FB06 2021-0051).

2.2. Materials

2.2.1. Explainer video and video script

Learning material in phase 1 was either an explainer video or the corresponding video-script about test anxiety. The content of the material was based on the content of the chapter "test anxiety" from the Handbook of Educational Psychology (Rost et al., 2018). It was structured in five sections: introduction, overview, definition of test anxiety, causes for test anxiety, and interventions. The introduction section was designed to trigger situational interest by starting with a relatable frame story and directly addressing participants' prior experiences (i.e. *"Perhaps you have experienced this before: you have an important exam approaching, […]"*) (e.g., Endres et al., 2020). Then, participants were given a brief overview as a structure to help them gain an initial understanding of the content (i.e. cognitive scaffold; Eitel & Scheiter, 2015). The overview was followed by the presentation of the learning content in everyday language, which also serves as an emotional design element. The learning material ended with a brief summary. The content for the first learning phase was selected to cover the main aspects of test anxiety in order to give a short overview regarding the definition of test anxiety, causes for test anxiety, and possible interventions.

The video-script comprised 814 words and was displayed on two pages (page 1: introduction and overview; page 2: learning content and summary). There was no time limit, but participants could not return to the previous page. The explainer video used sketched explanations. It was created based on the video-script using a video creator platform (https://simpleshow.com). The explainer video's audio-text was the video-script narrated by a female teacher student with minimal changes in the wording to fit the medium (for an exemplary screenshot, see Fig. 1). For the design of the explainer video, we adhered to the multimedia

Exemplary Screenshot from the Explainer Video.

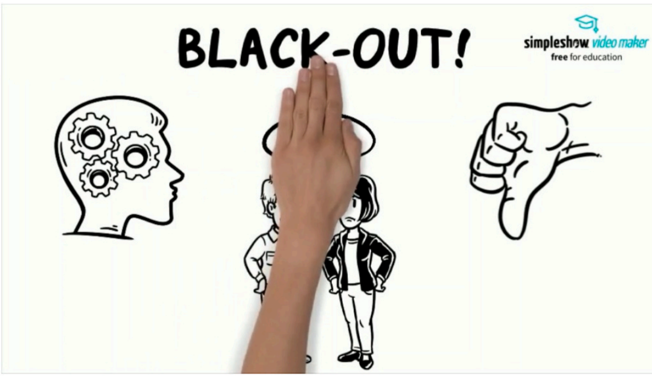

Fig. 1. Exemplary screenshot from the explainer video.

principles and design recommendations for videos (e.g., coherence principle, temporal contiguity; Mayer et al., 2020). Moreover, the video included social cues (e.g., a human hand that moved images in the video). Key words (e.g., cognitive interference) were presented in text form (Adesope & Nesbit, 2012). Participants could pause the video but not rewind it. Similar to the script, the video was divided into two parts and presented on two pages. Again, participants could not return to the previous page. The video duration was 5 min and 34 s.

The main difference between the script and the video was the wording, when the spoken/written text explicitly referred to the video/script. However, unfortunately, part of the summary section was also missing from the video in the study due to converting. Hence, the amount of information in the script and the video differed slightly in the last section. Importantly, the technical issue did not affect any additional information relevant to the prompts or the knowledge test.

2.2.2. Learning material in phase 2

For phase 2, participants received the chapter "test anxiety" by Cortina (2008) from the Handbook of Educational Psychology (Schneider & Hasselhorn, 2008). The ten content pages of the chapter were displayed on the same page so that participants could scroll through the text. Before the text was displayed, participants received a brief explanation for difficult terms (e.g., phylogenetic). There was no time limit for reading the chapter. The chapter was intended to complement the learning material from phase 1 by containing both the information from the video/script, as well as more in-depth information and additional aspects (e.g., psychological theories, empirical studies, individual differences). To be able to answer all post-test items correctly, learners had to learn the information from both learning materials.

2.2.3. Prompts

Based on Berthold et al. (2007), we created six prompts with five thereof relating to cognitive processes (i.e., retrieval/generative processes), and one to metacognitive processes. Within the three retrieval prompts, participants were asked to write down the definition, components, and symptoms of test anxiety (prompt-question 1: "*How do you define performance anxiety, what are its components and how do they affect learners? Please give your answer in 3-4 sentences.*"), to name different causes for test anxiety (prompt-questions 2: *"What are possible causes of severe test anxiety? Please give your answer in 1-2 sentences."*), and to describe possible interventions from the perspective of the affected person, the teacher, and the parents (prompt-question 3: "*How can learners, parents and teachers counteract test anxiety? Please give your answer in 3-4 sentences.*"). Within the two generative prompts, participants were asked to write down possible connections within the content (prompt-question 4: "*What connections between individual contents did you notice while working on the learning unit? Please give your answer in 1-2 sentences.*"), and to create an example of the most important content of the topic (prompt-question 5: *"Think of your own example of the most important content of the topic. Please give your answer in 2-3 sentences."*). To foster monitoring, we also asked participants to write down issues they found difficult to understand (prompt-question 6: *"What was difficult to understand? This will be followed by a second learning unit based on a textbook text on the topic (corresponds to the basic literature of the lecture). Please write down what you want to pay more attention to when reading the text (e.g. definition/causes/interventions)."*).

We displayed the prompts only after the first learning phase and not in between the learning material. This decision was based on the idea that the experimental design should resemble a real learning situation with a relatively short explainer video.

2.3. Measures

2.3.1. Prior knowledge

We assessed prior knowledge using one self-rating item and an open-ended question. First, we asked participants to indicate their current knowledge on the topic of test anxiety on a 0–100 percent scale (i.e. subjective prior knowledge). Then, we instructed them as follows: *"Please write down everything you know about the topic of test anxiety. Feel free to take a moment to think about it. Write at least 2–3 sentences."* Fifty percent of the participants' responses were evaluated by two independent raters in terms of the content's quantity and the quality (for double coded answers $ICC = 0.873$). For the double-coded responses, we then calculated an average score across both raters. For a total of five aspects (symptoms, definition, relevance, causes, and interventions), each answer could receive two to three points for quality, and one extra point per aspect for a detailed answer (i.e. quantity). Overall, participants could receive 0–12 points for quality and 0–5 points for quantity, resulting in a maximum total score for objective prior knowledge of 17 points.

2.3.2. Learning outcome measures

Overall, the post-test comprised 13 recall and transfer items in three different formats (i.e. multiple choice, cloze, and open-ended text).

Recall performance in this study describes learners' ability to answer questions that required information that could be directly derived from the learning material. We assessed recall performance with four items: One open-ended question, and three multiple choice items with 1–5 correct answer options each (e.g., "*Test anxiety is …*", example answer options: *"… the experienced fear in a certain exam situation"*, incorrect; "*… a personality trait*", correct). For the multiple-choice items, participants were given one point for each correct answer, resulting in a maximum score of 15 points. The open-ended question was the same as in the prior knowledge test and thus evaluated following the same criteria, resulting in a maximum score of 17 points. Similar to prior knowledge, 50% of the responses were double-coded by two independent raters (for double coded answers $ICC = 0.912$), and then the score of the double-coded responses was averaged across both raters. Overall, participants could score 0–32 points for recall performance (McDonalds' $\omega = 0.75$).

Transfer performance in this study describes learners' ability to answer questions that required a deeper understanding of the learning content. For instance, learners' had to apply the knowledge they gained in the learning phase to novel scenarios as well as making generalizations and/or inferences. We assessed transfer performance with five multiple-choice items, two open-ended questions, and two cloze items. Again, the multiple-choice items comprised 1–5 correct answer options each (e.g., *"Is it possible that the influence of performance anxiety is underestimated?"*, example answer options: *"No, because performance anxiety is compensated for."*, incorrect; *"Yes, because people with high performance anxiety might drop out of school earlier and thus systematically distort the results of long-term studies."*, correct). For each correct answer, participants received one point resulting in a maximum score of 25 points. In each of the two open-ended questions, an example situation was briefly described from everyday teaching that involved the test anxiety topic. To include all three intervention approaches, one of the questions focused on effective teacher behaviours, the other on effective parent behaviours and recommendations for those affected by test anxiety (for an example, see OSF). Participants could receive a maximum score of 10 points. Fifty percent of the responses to the open-ended questions were double coded by two independent raters (for double coded answers $ICC_{open\ transfer\ 1} = 0.763$ and $ICC_{open\ transfer\ 2} = 0.698$). For the double-coded answers, an average score was calculated for both raters. The first cloze question comprised four text gaps with 2–3 options and one correct answer each ("What are potential effects of high test anxiety?", e.g., *"Learners with high test anxiety usually invest* [*less*, incorrect]/[*more*, correct]/[t*he same amount*, incorrect] *of time in preparing for exams as other learners."*). The second cloze question comprised two text gaps with five answer options and one correct answer each. Because of a technical problem, only one of the two gaps could be scored. For each correct answer, participants received one point, resulting in a maximum score of 5 points. Overall, participants could score 0–40 points for transfer performance (McDonalds' ω = 0.52).

2.3.3. Prompts

We assessed participants' responses on the six prompts in terms of quantity and quality but used only the answers to the cognitive-related prompts (i.e. retrieval prompts and generative prompts) for the exploratory analyses. For quantity, we assessed the overall word count. To assess the quality, the responses to each of the six prompt-questions were coded, with 50% of responses double-coded ($ICC = 0.947$). For the three retrieval prompts, participants could receive a maximum total score of 25 points for correct responses. For the generative prompts, participants could receive 0–2 points for their own example, and one point for each mention of a relevant connection within the topic. We constructed a scale based on the five retrieval and generative prompts to measure possible effects of cognitive-related prompts (McDonalds' ω = 0.63). Participants could earn 0–2 points for the metacognitive prompt.

2.3.4. (Meta-)Cognitive measures

After the introduction, we asked participants to rate on a 0–100 percent scale the difficulty of learning with the rest of the script/video (i.e. EOL; ease of learning).

After each learning phase, we also assessed participants' judgement of learning (JOL), judgement of difficulty (JOD), mental effort, and cognitive load.

To assess participants' JOL, we asked them to rate on a 0–100 percent scale how confident they were that they would be able to answer questions about the learning content on a test at the end of the study. To calculate their metacognitive calibration (i.e., how well their self-assessment matched their objective learning outcomes) after the first learning phase, we first scaled their learning outcome on a 0–100 percent scale and then subtracted that value from their JOL-score from the first learning phase. For metacognitive calibration, a value over zero indicated overestimation and a value below zero underestimation. For participants in the prompt condition, we also assessed their specific judgements after the first learning phase on a 0–100 percent scale regarding the three content sections covered in the script/video. To assess participants' JOD, we asked them to rate on a 7-point Likert scale how difficult they found the learning unit to be (1 = *very easy* to 7 = *very difficult*).

We also assessed participant's active mental effort (e.g., "*I exerted myself to read the text*") and passive mental effort (e.g., *"It has been strenuous to read the text."*) with two items based on Klepsch and Seufert (2021). We also assessed *cognitive load* via the Cognitive Load Questionnaire (CLQ; Klepsch et al., 2017). For this, participants rated eight adapted statements regarding intrinsic (phase 1: Cronbach's α = 0.73, phase 2: Cronbach's α = 0.83), extraneous (phase 1: Cronbach's α = 0.80, phase 2: Cronbach's α = 0.79), and germane cognitive load (phase 1: Cronbach's α = 0.44; phase 2: Cronbach's α = 0.66). All items were assessed on a 7-point Likert scale (1 = *absolutely wrong* to 7 = *absolutely right*).

2.3.5. Situational interest

We assessed situational interest based on Endres et al. (2020) along two dimensions (i.e. affect and value; Schiefele, 2009). Participants' affect was measured by asking participants to rate the fit of the adjectives "exciting," "entertaining," and "boring" (reverse-coded) for the materials and content. Participants' value was measured by asking them to rate the fit of the adjectives "useful," "unnecessary" (reverse-coded), and "unimportant" (reverse-coded) for the materials and content. Both measures were rated on a 9-point Likert scale (1 = *not at all* to 9 = *very much*). Based on Endres et al. (2020), for triggered situational interest, we calculated a mean score for the items of the affect-towards-the-material subscales (phase 1: Cronbach's α = 0.80, phase 2: Cronbach's α = 0.65). For maintained situational interest, we calculated a mean score for the items of the affect-towards-the-content and value-towards-the-content subscales (phase 1: Cronbach's α = 0.78, phase 2: Cronbach's α = 0.80). Relying on Endres et al. (2020), we also calculated an overall measure of positive affect by calculating a mean score for the items of the affect-towards-the-material and affect-towards-the-content subscales for the first learning phase (phase 1: Cronbach's α = 0.80, phase 2: Cronbach's α = 0.74).

2.3.6. Further measures

Because the present study was conducted online, we included a short procedure check at the end of the experiment. First, all participants were asked to answer truthfully, as this would not put them at any disadvantage. We then assessed their genuine participation with three items (e.g., "*Did you use external resources (*e.g.*, Google, notes, etc.) to help you complete the post-test?*" "*yes*" vs. "*no*"), followed by a question regarding which device they had used to participate in the study to check for technical issues.

The measures below were also assessed for exploratory reasons but not included in our analyses: Personal interest (Schiefele et al., 1993), academic self-concept (Schöne et al., 2002), learning strategies (Klingsieck, 2018), intrinsic motivation (Wilde et al., 2009), and prior experience with explainer videos.

2.4. Procedure

After having provided informed consent to participate in the online study, participants answered demographic questions, questions on their personal interest and their academic self-concept regarding educational psychology, as well as their general use of learning strategies and their prior knowledge on the topic of test anxiety. They were then assigned to either the video or the video-script condition for the first learning phase.

After reading or watching the short introduction, participants rated their perceived difficulty of learning with the rest of the script or video (i.e. ease of learning). Then, participants read or watched the rest of the video-script or video. They then rated their cognitive load, perceived difficulty, and situational interest regarding the content and design of the learning material. Afterwards, half of the participants received six prompts, and then rated how well they would be able to answer questions about the topic in a test at the end of the study. Participants in the no-prompt condition rated their judgement of learning only globally, while participants in the prompt condition also rated their specific judgements regarding the three content sections covered in the script or video.

Then, the second learning phase started. There was no time limit for reading the book chapter, so that all participants were able to finish the learning phase at their own pace. After the second learning phase, we again assessed participants' judgement of learning, cognitive load, judgement of difficulty, and situational interest. This was followed by the post-test. Because the study was conducted as an online study, we were not able to control their potential use of learning aids, thus we explicitly instructed them to put those aside.

After participants completed the post-test at their own pace, we assessed their interest and intrinsic motivation regarding their entire learning experience. At the end of the study, we asked them to answer three procedure-check items addressing their faithful participation, and to indicate which device they used in the study. Finally, all participants were debriefed and forwarded to a new questionnaire where they could enter their e-mail addresses if they wanted to participate in the lottery for the gift cards.

3. Results

For all statistical analyses, we applied an alpha level of $\alpha = 0.05$, and one-tailed test statistics for directional hypotheses. As indices for effect size, for Analyses of Variances and Covariance we used η_p^2 with values of 0.01, 0.06, 0.14, and for t-tests Cohens' d with values of 0.2, 0.5, 0.8 as small, medium, and large effect size, respectively. For mediation analyses, we used dummy coding for the experimental conditions medium (coding [0] video-script and [1] video) and prompts (coding [0] no prompts and [1] prompts).

3.1. Preliminary analyses

Analyses of Variance (ANOVAs) revealed no significant differences between experimental conditions regarding subjective, $F < 1$, or objective prior knowledge, $F(1,129) = 1.62$, $p = .206$, $\eta_p^2 = 0.01$. Only objective prior knowledge correlated significantly with post-test performance ($r = 0.20$, $p = .021$). We thus included objective prior knowledge as a covariate in Analyses of Covariance (ANCOVAs) and mediation analyses with learning outcome as dependent variable.

3.2. Main analyses

3.2.1. Situational Interest Hypothesis (H1)

We conducted a one-tailed independent-samples *t*-test to compare participants' triggered situational interest after the first learning phase between the video and video-script condition. In line with H1a, our results revealed a significant difference: Learners in the video condition reported significantly higher triggered situational interest (phase 1) than learners in the video-script condition, $t(131) = -4.69$, $p < .001$, *Cohens' d* = 0.82 (see Table 1 for the descriptive statistics).

To test whether triggered situational interest would result in higher maintained situational interest and in learners investing more learning time in the second learning phase, and thus performing better at the post-test (H1b), we conducted two separate three-step-mediation analyses for recall and transfer performance, respectively (Model 6, Hayes, 2022). For each of the two dependent variables, we entered medium (coding [0] video-script and [1] video) as independent variable, triggered situational interest (phase 1), maintained situational interest (phase 2) and invested learning time (phase 2) as mediators, and prior knowledge as covariate. Against our expectations, our results indicated no significant specific indirect effect either for recall, $b = 0.02$, $SE = 0.03$, 95% CI [−0.02 | 0.10] (partially standardized indirect effect: $b = 0.003$, $SE = 0.01$, [-0.004 | 0.02]) or transfer performance, $b = 0.01$, $SE = 0.02$, 95% CI [−0.01 | 0.07] (partially standardized specific indirect effect: $b = 0.004$, $SE = 0.01$, [−0.004 | 0.02]).

3.2.2. Illusion of understanding hypothesis (H2)

Because lower mental effort and a lower judgement of difficulty for the first learning phase can reflect an illusion of understanding, we investigated whether there would be a difference between the video and the video-script condition. For this, we conducted three one-tailed independent-samples t-tests with active mental effort (phase 1), passive mental effort (phase 1) and perceived difficulty of the learning content (phase 1) as dependent variables (see Table 1 for the descriptive statistics). In line with our expectations, learners in the video condition reported significantly *less* active mental effort, $t(131) = 1.86$, $p = .032$, *Cohens' d* = 0.33, *less* passive mental effort, $t(103.7) = 1.77$, $p = .040$, *Cohens' d* = 0.32, and assessed the content overall as being *less* difficult, $t(131) = 1.67$, $p = .049$, *Cohens' d* = 0.29, compared to learners in the video-script condition, thus supporting the assumption that explainer videos seem easier than text (H2a).

Table 1
Descriptive statistics of important variables in phase 1.

	Condition				Range
	Explainer video		Video-script		
	n = 76		n = 57		
	M (SD)		*M* (SD)		
Triggered situational interest (phase 1)	5.82 (1.65)		4.43 (1.77)		1–9
Maintained situational interest (phase 1)	7.27 (1.08)		6.98 (1.10)		1–9
Active mental effort (phase 1)	3.50 (1.53)		4.00 (1.54)		1–7
Passive mental effort (phase 1)	2.28 (1.18)		2.70 (1.50)		1–7
Judgement of difficulty (phase 1)	2.70 (0.86)		2.98 (1.11)		1–7
	prompts	no prompts	prompts	no prompts	
	n = 30	n = 43	n = 22	n = 35	
	M (SD)	*M* (SD)	*M* (SD)	*M* (SD)	
Cognitive prompts (quality)	11.43 (4.00)	–	9.07 (2.86)	–	0-27 (min.)*
Cognitive prompts (amount of words)	185.00 (84.98)	–	163.14 (77.02)	–	–
Metacognitive prompt (quality)	1.10 (0.89)	–	1.18 (0.66)	–	0–2
Metacognitive prompt (amount of words)	11.43 (12.55)	–	13.36 (13.45)	–	–
Metacognitive calibration (phase 1)	−11.41 (25.83)	−8.66 (19.03)	−4.70 (22.50)	−3.35 (18.29)	−100-100

Note: For the second generative prompt, participants could receive one point for each mention of a relevant connection. Thus, there was no upper point-limit for this question.

We also expected that without prompts learners watching the explainer video would tend to overestimate their knowledge level (H2b). To test this hypothesis, we conducted a one-tailed independent simple *t*-test with medium (video vs. video-script) as independent variable and metacognitive calibration as dependent variable for the no-prompt condition. Against our expectations, for learners without prompts, we identified no significant difference between the video and video-script conditions, $t(76) = 1.25$, $p = .108$, *Cohens' d* = 0.28.

Moreover, we expected that an illusion of understanding in the first learning phase would lead to less learning time and effort in the second learning phase, thus resulting in poorer learning outcomes (H2c). Thus, for the no-prompt condition, we conducted two separate mediation analyses with the dummy coded medium condition ([0] video-script and [1] video) as independent variable, the two sequential mediators active mental effort and learning time in phase 2, for recall and transfer performance as dependent variables, respectively, and prior knowledge as covariate (Model 6; Hayes, 2022). Against our hypotheses, our results indicated no significant specific indirect effects either for recall, $b = -0.21$, $SD = 0.18$, 95% CI [−0.66 | 0.05] (partially standardized indirect effect: $b = -0.05$, $SE = 0.04$, [−0.15 | 0.01]) or transfer performance, $b = -0.18$, $SE = 0.15$, 95% CI [−0.52| 0.05] (partially standardized indirect effect: $b = -0.06$, $SE = 0.05$, [−0.16 | 0.02]).

3.2.3. Prompts support learning process hypothesis (H3)

To test our hypothesis that prompts (vs. no prompts) after the first learning phase would result in more accurate metacognitive judgements about the learning outcomes in a post-test (H3a), we calculated an ANOVA with metacognitive calibration (Phase 1) as dependent variable. Due to missing values, three participants were excluded from the analysis. Against our expectation, there was neither a main effect of medium, $F(1,125) = 2.33$, $p = .129$, $\eta_p^2 = 0.02$, nor of prompts, $F < 1$, nor a significant interaction effect between medium and prompts, $F < 1$.

Moreover, we expected that learners in the prompt condition (vs. in the no-prompt condition) would invest more mental effort (H3b) and more learning time (H3c) in phase 2. Therefore, we conducted two separate ANOVAs with active mental effort (phase 2) and learning time in the second learning phase as dependent variables.

Our results for active mental effort (phase 2) indicated a significant main effect for the prompt condition, $F(1,129) = 6.57$, $p = .012$, $\eta_p^2 = 0.05$, but not for the medium condition, $F(1,129) = 2.02$, $p = .158$, $\eta_p^2 = 0.02$. Against H3a, our results indicated that with prompts learners reported *less* active mental effort during the second learning phase. However, based on the significant interaction between the experimental conditions, $F(1,129) = 10.27$, $p = .002$, $\eta_p^2 = 0.07$, this difference is mainly attributable to the significant difference in the video-script condition (see Table 2 for the descriptive statistics). Results from simple comparisons indicated that learners with prompts in the video-script condition reported significantly less active mental effort than those without prompts ($M_{diff} = -1.71$, $SE = 0.45$, $p < .001$). For the video condition, this difference was not significant ($M_{diff} = 0.19$, $SE = 0.39$, $p = .624$). Moreover, with prompts, learners in the video condition reported significantly more active mental effort compared to learners in the video-script condition ($M_{diff} = 1.37$, $SE = 0.46$, $p = .004$). For the condition without prompts, the difference was not significant ($M_{diff} = 0.53$, $SE = 0.37$, $p = .156$).

Regarding invested learning time in phase 2, the medium (video vs. script) also revealed no main effect, $F < 1$. However, we observed a significant main effect for the prompt condition, $F(1,124) = 4.05$, $p = .046$, $\eta_p^2 = 0.03$. But against H3b, learners with prompts invested significantly *less* learning time than those without. Furthermore, the interaction between experimental conditions was not significant, $F(1,124) = 3.56$, $p = .061$, $\eta_p^2 = 0.03$.

Table 2
Descriptive statistics of important variables in phase 2.

	Condition				Range
	Explainer Video		Video-script		
	prompts	no prompts	prompts	no prompts	
	n = 30	n = 43	n = 22	n = 35	
	M (SD)	*M* (SD)	*M* (SD)	*M* (SD)	
Triggered situational interest (phase 2)	2.82 (0.98)	2.96 (1.69)	3.59 (1.76)	3.50 (1.60)	1–9
Maintained situational interest (phase 2)	5.79 (1.38)	6.00 (1.26)	5.95 (1.48)	6.36 (1.28)	1–9
Active mental effort (phase 2)	5.23 (1.72)	5.04 (1.61)	3.86 (1.83)	5.57 (1.52)	1–7
Passive mental effort (phase 2)	1.97 (0.85)	2.48 (1.33)	2.77 (1.63)	2.66 (1.43)	1–7
Judgement of difficulty (phase 2)	5.00 (0.89)	4.70 (1.07)	5.09 (0.92)	4.63 (1.09)	1–7
Recall performance	17.58 (5.12)	14.54 (4.27)	13.14 (3.54)	16.37 (4.71)	0–32
Transfer performance	30.40 (3.09)	30.23 (3.03)	28.93 (2.35)	30.99 (3.25)	0–40
Metacognitive calibration (phase 2)	−31.51 (20.00)	−26.88 (23.40)	−26.20 (18.20)	−23.43 (14.46)	−100-100
Intrinsic motivation	3.39 (0.73)	3.27 (0.68)	2.80 (0.65)	3.16 (0.62)	

A possible explanation for why we failed to detect a positive effect of prompts on mental effort and learning time in the second learning phase might be that not all learners needed to use the indirect feedback from answering the prompts for the second learning phase. Possibly, only learners who accurately judged their performance in the retrieval prompts as poor used this information to adjust their behaviour in the second learning phase. Against this background, we performed two exploratory quadratic regression analyses for those learners whose performance in the retrieval prompts were overall below 50% percent. In the first quadratic regression analysis, we entered metacognitive calibration (phase 1) for the retrieval prompts as independent variable and invested learning time (phase 2) as dependent variable. Against our assumption, the regression model with the quadratic term was not significant, $F(2,36) = 2.04$, $p = .144$, $\eta_p^2 = 0.05$. In the second quadratic regression analysis, we entered metacognitive calibration for the retrieval prompts as independent variable, and active mental effort (phase 2) as dependent variable. Again, however, the regression model with the quadratic term was not significant, $F < 1$.

Another possible explanation is that learners in the video-script condition with prompts were demotivated, and thus invested less mental effort and learning time in the second learning phase. To test this assumption, we calculated an exploratory ANOVA with participants' intrinsic motivation as dependent variable, and medium (video vs. video-script) as well as prompt condition (prompt vs. no-prompt) as independent variables. Results revealed a significant main effect of medium, $F(1,129) = 8.52$, $p = .004$, $\eta_p^2 = 0.06$, indicating a higher intrinsic motivation for learners in the video condition. For the prompt condition, the main effect was not significant, $F < 1$. The interaction between medium and prompt condition was also not significant, $F(1,129) = 3.87$, $p = .051$, $\eta_p^2 = 0.03$. Descriptively, however, learners in the video-script condition with prompts reported the lowest intrinsic motivation regarding their learning experience at the end of the study (see Table 2 for mean scores).

It is possible that the demotivation for learners in the video-script condition with prompts was related to more effort in answering the prompts after the first learning phase which might have let to more motivational depletion during the second learning phase. To test this assumption, we calculated two exploratory independent t-tests. First, we entered the time learners spent on the prompts as dependent variable and medium (video vs. video-script) as independent variable. There was no significant difference between the video and the video-script condition, $t(50) = -0.16$, $p = .870$, *Cohens' d* $= 0.05$, indicating that learners in both conditions spent equally long on answering the prompt questions after the first learning phase. Second, we entered the number of words participants wrote to answer the prompt questions as dependent variable and medium (video vs. video-script) as independent variable. Again, the difference was not significant, $t(50) = -0.82$, $p = .417$, *Cohens' d* $= 0.23$. On average, learners in both conditions wrote a similar number of words to answer the prompt questions. We also tested whether there would be a correlation between time spent on answering the prompts and time spent in the second learning phase. In contrast to the motivational depletion assumption, results indicated a positive correlation, $r = 0.405$, $p = .004$, indicating that learners who spent more time to answer the prompt questions also invested more learning time in the second learning phase.

3.2.4. *Prompts support learning outcomes hypothesis (H4)*

To test whether the medium (video vs. video-script) and prompt condition interaction (prompt vs. no prompt) would influence the learning outcome, we conducted two ANCOVAs for recall and transfer with prior knowledge as covariate (see Table 2 for the descriptive statistics).

For *recall*, there was neither a main effect for medium, $F(1,128) = 2.40$, $p = .124$, $\eta_p^2 = 0.018$, nor for the prompt condition, $F < 1$. However, our results indicated a significant interaction between experimental conditions, $F(1,128) = 13.57$, $p < .001$, $\eta_p^2 = 0.10$. Simple comparisons within the prompt condition revealed a significant difference between the video and video-script condition ($M_{diff} = 4.18$, $SE = 1.25$, $p = .001$). In line with H4b, learners in the video condition performed better than those in the script condition. With regard to the no-prompt condition, simple comparisons revealed no significant difference ($M_{diff} = -1.72$, $SE = 1.00$, $p = 0.087$). Furthermore, simple comparisons within the video-script condition revealed a significant effect of prompts ($M_{diff} = -2.88$, $SE = 1.22$, $p = .019$), with learners in the prompt condition performing *worse* than those in the no-prompt condition. Simple comparisons within the video condition also revealed a significant effect of prompts ($M_{diff} = 3.01$, $SE = 1.04$, $p = .004$). Here, learners in the prompt condition performed significantly *better* than those in the no-prompt condition (Fig. 2).

Against H4, for *transfer*, there was neither a main effect of medium, $F < 1$, nor of the prompt condition, $F(1,128) = 2.68$ $p = .104$, $\eta_p^2 = 0.02$, nor a significant interaction, $F(1,128) = 3.74$, $p = .055$, $\eta_p^2 = 0.03$.

3.3. Exploratory analyses

To test whether explainer videos serve as cognitive scaffold, we tested whether learners who watched the video (vs. read the video-script) would be better at answering the retrieval and generative prompt questions after the first learning phase. For this, we calculated a mean score of the quality scores for the answers to the retrieval and the generative prompt questions. An ANCOVA for learners in the prompt condition with medium (video vs. video-script) as independent variable and prior knowledge as covariate, was not significant, $F(1,49) = 4.00$, $p = .051$, $\eta_p^2 = 0.08$. Descriptively, however, learners in the video condition wrote better answers to the retrieval and generative prompts than those in the video-script condition. Moreover, we tested whether learners in the video condition spent more time answering the prompt-questions. Results of an independent-samples *t*-test did not show significant differences between the video and video-script condition, $t(50) = -0.16$, $p = .871$, *Cohens' d* $= 0.46$.

Furthermore, we conducted two exploratory mediation analyses for learners in the prompt condition (Fig. 3). We entered the dummy coded medium condition (coding [0] video-script and [1] video) as independent variable, prior knowledge as covariate, time-on-prompts and quality of prompt-answers for the retrieval and generative prompts as potential mediators, and recall and transfer performance as dependent variables, respectively (Model 4; Hayes, 2022).

Results of the mediation analysis with recall as dependent variable indicated that only the quality of the prompt-answers partially mediates the effect of medium (video vs. video-script) on recall, $b = 1.91$, $SE = 0.90$, 95% CI [0.24 | 3.80] (partially standardized indirect effect: $b = 0.38$, $SE = 0.16$, 95% CI [0.05 | 0.71]), but not time-on-prompts. For transfer as dependent variable, we found similar results. Again, only the quality of the prompt-answers mediates the effect of medium (video vs. video-script) on recall, $b = 0.79$, $SE = 0.48$, 95% CI [0.03 | 0.87] (partially standardized indirect effect: $b = 0.27$, $SE = 0.16$, 95% CI [0.12 | 0.62]).

Recall Performance as a Function of Condition

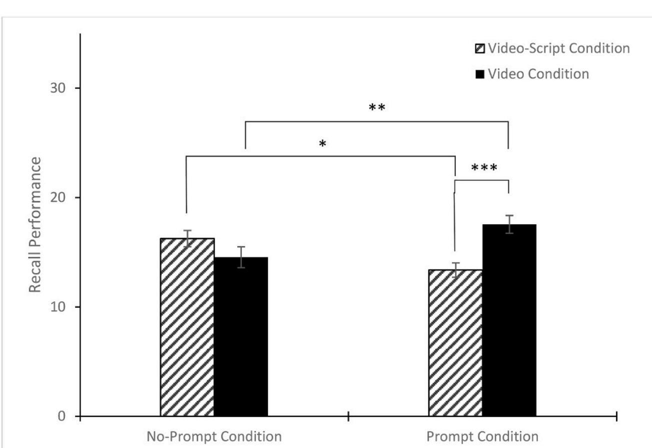

Fig. 2. Recall performance as a function of condition.

4. Discussion

The aim of the present study was twofold. First, we investigated whether watching explainer videos before text fosters an illusion of understanding and is thus detrimental for subsequent learning with textbook material. Second, we investigated whether prompts provided as content-related and monitoring-related questions would prevent learners from developing an illusion of understanding, and thus foster learning with explainer videos and subsequent textbook material.

In summary, although we found no evidence that learners in our study developed an illusion of understanding after watching the explainer video, they did benefit from the prompt-questions for their learning outcome. Exploratory analyses indicated that this beneficial effect was mainly driven by the quality of the prompt-answers. Surprisingly, learners in the video-script condition did not benefit from prompt-questions, and actually performed worse in the post-test than those without.

Regarding the situational interest hypothesis, in line with our expectations, we found that learners in the video condition reported more triggered situational interest towards the learning material and the content than learners in the video-script condition (H1a). This concurs with previous research showing that explainer videos can trigger situational interest (Endres et al., 2020) - a precondition for further learning processes. Against the situational interest hypothesis, however, learners failed to invest more learning time after watching the explainer video. Thus, the explainer video was helpful to catch learners' attention, but that had no beneficial effects on the subsequent learning with the textbook chapter.

As a competing hypothesis, we formulated the illusion-of-understanding hypothesis (H2). We expected that if explainer videos fostered an illusion of understanding (Paik & Schraw, 2013), learners in the video condition would exhibit signs of superficial processing, (e.g., perceiving the content to be easier, overestimating their level of understanding), and consequently invest less effort and time in the second learning phase. In line with our expectations, we found that learners in the video condition did report significantly less active and passive mental effort, and assessed the content to be less difficult, which might be initial indications that those learners invested less effort in processing the content in the explainer video.

Nevertheless, that had no significant effect on their subsequent learning. Our results do not reveal that learners overestimated their knowledge after having watched the explainer video, or that this resulted in more ineffective learning in the second learning phase and thus in poorer learning outcomes. Against our expectations, we did not find that learners in the video condition were significantly less well metacognitively calibrated compared to learners in the video-script condition. Similarly, Tarchi et al. (2021) found that even if learners in a video condition perceived higher self-efficacy regarding their learning compared to learners in a text condition, this was not mirrored by their calibration judgements. Interestingly, in our study, learners in the video condition tended to underestimate their performance - and not overestimate it, as we would have expected. Moreover, providing prompts descriptively increased the underestimation. Consequently, our results for metacognitive calibration did not support the illusion of understanding hypothesis (H2).

However, our results showed that learners in the video condition reported more passive cognitive processing during the first learning phase compared to learners in the video-script condition. It is possible

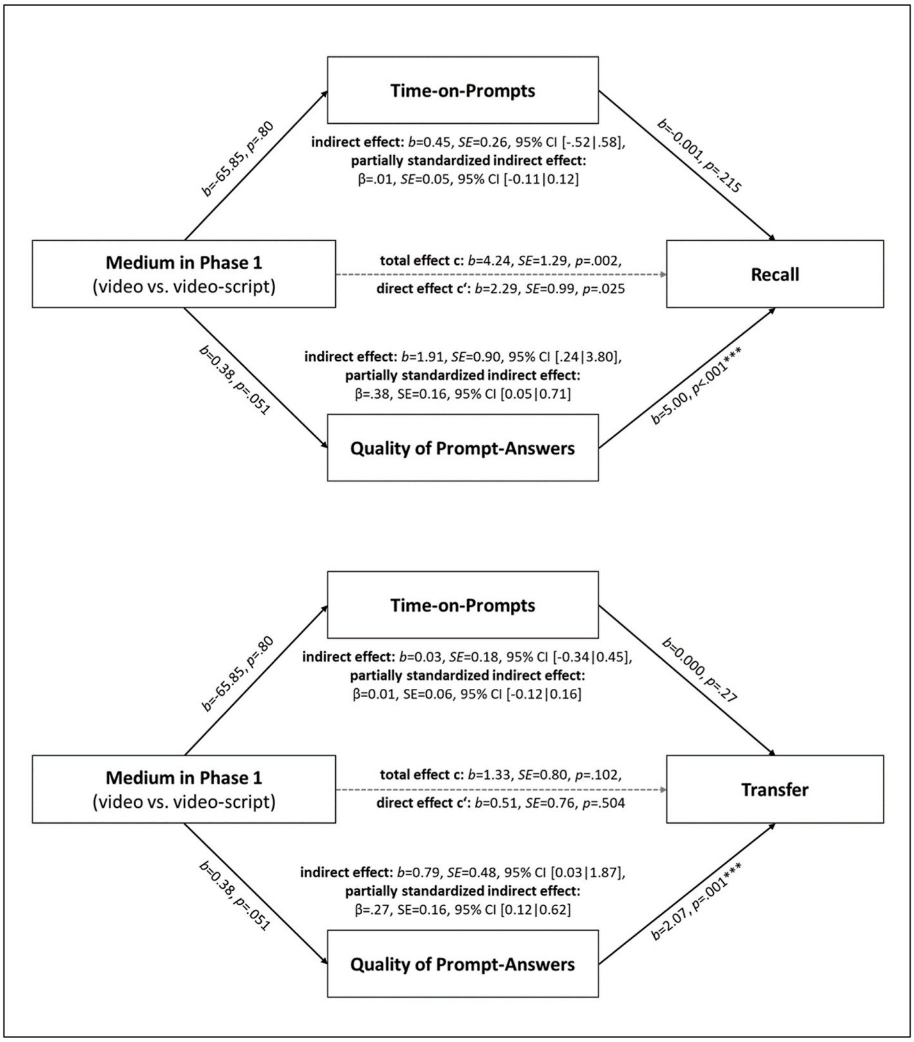

Fig. 3. Results of the exploratory mediation analyses with recall and transfer as dependent variables.

that they became aware that this might be problematic for learning, which might have resulted in a more humble metacognitive mindset. This might also explain the fact that learners in the video condition with prompt-questions descriptively underestimated their knowledge level even more than those in the video condition without prompt-questions. This is because if learners became aware of their superficial processing after watching the explainer video and were also prompted to explicitly reflect on their learning process by answering the prompt-questions, this could have led to a greater underestimation, as learners might have also compensated for the perceived easiness of the video. In sum, it is possible that learners experienced an illusion of understanding during the first learning phase and reflecting on their learning behaviour might have resulted in learners underestimating themselves more.

Overall, even though we found that learners watching the explainer video were more prone to superficial processing which might indicated an illusion of understanding during watching the video, this was not reflected in their metacognitive calibration afterwards. Hence, our results do not support the illusion-of-understanding hypothesis (Salomon, 1984).

Concerning the use of prompt-questions after the first learning phase, we found - in contrast to previous research (e.g., Berthold et al., 2007) - that providing prompts was not necessarily helpful. In contrast to our expectations that providing learners with prompt-questions would help them monitor and consolidate their knowledge (e.g., Eitel et al., 2022; Roelle et al., 2022), we failed to detect any general beneficial effect of prompts on learners' metacognitive calibration (H3) or learning outcomes (H4).

A possible explanation for why we failed to detect a general positive effect of prompt-questions on mental effort (H3b) and learning time in the second learning phase (H3c) is that this might only apply for learners who correctly rated their performance as poor after the first learning phase. In contrast, learners who correctly rated their performance as good might have required less effort and time because they could skim information they already knew. Therefore, we performed an exploratory non-linear regression analysis for learners with a recall performance below 50% to test for a U-shaped relationship between learners' metacognitive calibration based on their recall performance and mental effort or learning time in the second learning phase. Unfortunately, our data did not support this alternative explanation.

For learning outcomes, our analyses' results indicate that while learners in the video condition (in line with our expectations and previous research, e.g., Szpunar et al., 2014) benefitted from prompts regarding their learning outcome, learners in the video-script condition even performed worse than those without prompts. It is possible that learners in the video-script condition with prompts were more demotivated compared to the learners in the other conditions. In line with this assumption, we found that learners in the video-script condition with prompts reported the lowest intrinsic motivation at the end of the study. We tested in exploratory analyses whether an increased effort in form of invested time and produced content in the prompt phase might explain a 'motivational depletion' for learners in the script condition with prompts. In contrast to this possible explanation, however, we did not find significant differences for invested time and amount of content regarding the prompts between learners in the video-script and the video condition. Rather, we found a positive correlation between the time learners spent on the prompts and the time learners invested in the second learning phase, which also speaks against the assumption that learners who invested more time to answer the prompt-questions were less motivated to invest learning time in the second learning phase. Overall, there appears to be a motivational problem in the video-script condition with prompts but we were not able to determine the source of the problem. It might be interesting to take a closer look at a possible influence of epistemically-related emotions in this context (Pekrun et al., 2017) in future studies.

For learners in the video condition with prompts, exploratory analyses revealed that the superiority of the video compared to the video-script was mainly driven by the quality of the prompt-answers. Learners in the explainer video condition delivered descriptively higher quality answers for the retrieval and generative prompts which in turn related positively to better recall performance.

Due to our study design, time-on-task in phase 1 (i.e. time spent on video/video-script and time spent on answering the prompt-questions) differed between the prompt conditions (prompt vs. no-prompt), which might have influenced the effects of prompts on learning outcomes in our study. However, answer quality rather than the mere time learners took to answer prompt-questions led to better recall performance in the video condition. In the conditions with prompts, learners spent a similar amount of time answering the prompt-questions. Yet, learners in the video condition were better at answering the prompt-questions, which in turn lead to better recall performance. In contrast, learners in the video-script condition invested a similar amount of time in answering the prompt-questions. Consequently, we would argue that they were similarly motivated to answer the prompt-questions. However, they performed worse in answering the prompt-questions after phase 1 and in answering the recall questions at the end of the study. Taken together, these results indicate that learners in the video condition were better able to utilise the prompt-questions to their advantage.

Against this backdrop, explainer videos in combination with prompt-questions can have a positive influence on subsequent learning with textbook material. On the one hand, explainer videos can trigger situational interest in the topic (Endres et al., 2020), and on the other hand, they can – like pictures – scaffold subsequent learning from text (Eitel & Scheiter, 2015). However, our results also show that the positive effects of explainer videos only emerged when prompt-questions were provided. Although we detected no empirical evidence that learners in our study benefitted from the prompt-questions in metacognitive monitoring and regulation terms, it still appears that learners in the video condition were able to use the prompts to consolidate the knowledge they had acquired in the explainer video, and thus benefitted from prompts. Overall, in terms of the use of prompts, our results reveal that the effective provision of instructional support is often not as simple as it might first appear.

4.1. Limitations, implications and conclusions

We conducted the present study as an online experiment. On the one hand, this can be problematic, as we had no control over what the learners actually did during the experiment. To counter this problem, we implemented control questions at the end of the study, checked for a reasonable participation timeframe, and excluded participants who stated that they were too distracted during learning and used external resources for the post-test. Another challenge associated with the reduced control is the limited insight into the reasons why participants withdrew from the study. Overall, the dropout was in an acceptable range for an online learning study. After the start of the first learning phase, 40 participants terminated the study early. While the dropout of these 40 participants was not equally distributed across the four conditions ($X^2(3) = 20.01$, $p < .001$), a post-hoc analysis revealed that this was mainly due to the fact that overall learners in the explainer video condition without prompts dropped out significantly less than expected. A closer look at the dropout revealed that in the first learning phase significantly more learners from the video condition ($n = 12$) compared to learners from the video-script condition ($n = 3$) dropped out ($X^2(1) = 3.14$, $p = .041$). It is possible that they dropped out due to smaller technical issues (e.g., problems with video sound). Besides smaller technical issues, it is also possible that some participants in the video condition did not like the design of the video (e.g., the pictures, the voice, the speed), and decided to withdraw because it was not possible to skip the video or to increase the speed. However, it is also possible that individual characteristics affected the dropout. Consequently, we took a closer look at potential variables that might have affected a dropout. For this, we conducted Mann-Whitney-U-Tests between participants who withdrew from the study in the first learning phase and the other participants. We neither found significant differences for personal interest in educational psychology ($U = 1537.000$, $Z = -0.002$, $p = .998$), nor academic self-concept ($U = 1711.000$, $Z = 0.912$, $p = .362$), nor self-rated prior knowledge ($U = 1772.000$, $Z = 1.210$, $p = .226$). During answering the prompt-questions, an equal number of participants dropped out from the video condition ($n = 12$) and the video-script condition ($n = 11$). In the second learning phase, the dropout ($n = 9$) was equally distributed between the experimental conditions ($X^2(3) = 4.49$, $p = .213$).

An advantage of the online setting, on the other hand, is that it enables greater ecological validity, as learners – similar to a real learning situation – self-regulated their learning with the learning material in a familiar environment. Of course, it must be taken into account that the learning material used in the study is to a certain extent artificial. For example, the use of the video-script as learning material is not entirely comparable with more complex text materials used by learners in natural learning contexts. Nonetheless, the video-script presents a narrative summary of learning-relevant information which can also be found in a similar form in textbooks. However, for further research, it would be interesting to replicate the study in a lab and to collect more objective process measures, such as eye-tracking data, to see what learners actually do while learning with different materials. Moreover, it could also be interesting to examine factors that might influence the effectiveness of explainer videos in authentic learning situations, such as prior knowledge, mind wandering, distraction, and the use of supplemental resources during test taking in more depth.

With regard to transfer performance as learning outcome measure, it is possible that we failed to detect significant results because the learning phases were too short. It is also possible that an effect of the experimental conditions on transfer performance would only be visible in a delayed post-test. Nonetheless, in typical studies on (multimedia) learning, study sessions are even shorter than it was here, and Mayer and colleagues often find solid effects on transfer outcomes (see e.g., Mayer & Fiorella, 2022; Noetel et al., 2022, for a meta-meta-analysis). It would be interesting to investigate possible effects of medium and prompts on transfer performance for longer learning phases as well as with delayed testing.

Another limitation of our study has to do with the fact that we employed a well-designed explainer video. Our results could have differed, had we employed less well-developed explainer videos, especially in terms of metacognitive calibration, as they might prompt learners to overestimate their knowledge more. Our study participants rated the explainer video as being more interesting than the text only version (i.e. the video-script), but the mean value was still rather close to the middle area of the scale. Hence, it is possible that the explainer video was not interesting enough to foster overestimation. Furthermore, we only used text material (i.e. textbook chapter) in the second learning phase. It is possible that effects would change for other types of materials such as video-lectures because the explainer video's positive effect on learning outcomes in the prompt-condition might be partially due to a multimedia effect (e.g., Mayer, 2022). In contrast to previous research (e.g., Bai et al., 2022; Moos & Bonde, 2016; Zheng et al., 2023), we decided to not include the prompts in the video but displayed them only after the learners watched the whole explainer video. There is empirical evidence that for longer videos it is beneficial to include prompts in the video (e.g., Moos & Bonde, 2016; Zheng et al., 2023). Thus, it might be interesting to investigate whether there would also be beneficial effects for prompts on subsequent learning processes if they were included in the explainer video.

Another limitation concerns the measure we used to assess metacognitive calibration. We calculated the measure using the post-test performance at the end of the study and the judgement of learning (JOL) after the first learning phase. For one, the measure we used in the study reflects a convincement regarding participants' later performance on a global level but was not based on the percentage of correctly answered questions. However, the merged measure of post-test performance and judgement of learning in the form we used reflects the deviation from the learners' conviction about their performance after the first learning phase and their actual performance (global prediction; Baars et al., 2020). It might be interesting to use a more specific measure for learners' JOL in future studies. For another, we administered the JOL measure in the prompt-condition only after the prompt questions, and not before and after answering the questions. Therefore, we cannot investigate a possible change in the assessment due to answering the prompt questions within a learner. Based on our study design, we can only compare the respective JOL assessments between conditions (prompt vs. no-prompt). It might be interesting to take a closer look at a potential change of the JOL assessment in relation to the performance in prompt questions in future studies.

Another problem regarding the monitoring calibration measure relates to the test expectancy of the participants. Participants did not receive specific information on what kind of questions they had to expect in the post-test when answering the judgement of learning items after the first learning phase. It is therefore not entirely clear what the participants based their judgement on. Possibly, participants in the prompt condition expected open-ended questions based on the format of the prompt questions. However, participants in the no-prompt condition had no prior anchor of test items for their judgement. Accordingly, they might have had more difficulties in judging their performance accurately because they had less test information compared to participants in the prompt condition. Previous research, however, indicates that the knowledge about test demands alone does not automatically increase monitoring accuracy (Eitel, 2016). Further research should consider this, however, and give participants more information about the post-test when asking them to judge their later performance.

Against this background, further research is needed to be able to generalise our findings to other types of explainer videos and learning materials.

Overall, our findings contribute to research evidence on using explainer videos for formal learning purposes by showing that prompt-questions can foster learning with explainer videos. In line with previous research, we also found that explainer videos can induce situational interest, but in contrast to other research, we found no evidence for the assumption that explainer videos foster an illusion of understanding. Interestingly, we failed to observe that prompts exert a generally beneficial effect – we even found that the video-script condition prompts hampered learning. Concerning practical recommendations, it is important to note that more time is needed when prompts are added to the instruction, which is an issue when there are close time restrictions (e.g., school hours). In a nutshell, our results imply that the appropriate provision of instructional support such as prompts is often not as simple as it might first appear.

CRediT authorship contribution statement

Marie-Christin Krebs: Writing – review & editing, Writing – original draft, Formal analysis, Conceptualization. **Katharina Braschoß:** Investigation, Conceptualization. **Alexander Eitel:** Writing – review & editing, Writing – original draft, Formal analysis, Conceptualization.

Acknowledgements

We thank Charlotte Hild for her support in data collection.

Appendix A

Table 3
Time on Learning Task in Phase 1 and Phase 2

	Condition			
	Explainer video		Video-script	
	prompts	no prompts	prompts	no prompts
	n = 30	n = 46	n = 22	n = 35
	M (SE)	*M* (SE)	*M* (SE)	*M* (SE)
Time spent on material in phase 1 (sec)[a]	448.47 (29.88)	466.72 (24.95)	196.14 (35.71)	276.23 (27.66)
Time spent on prompts in phase 1 (sec)[b]	1042.27 (176.80)	–	990.95 (211.31)	–
Time spent on material in phase 2 (sec)	678.97(138.07)	694.02(96.78)	443.90(169.10)	910.58(131.65)

Note.
[a] Two participants were excluded based on an exploratory data analysis due to extreme processing times for learning material 1.
[b] Two participants were excluded based on an exploratory data analysis due to extreme times for answering the prompt-questions.

References

Ackerman, R., & Leiser, D. (2014). The effect of concrete supplements on metacognitive regulation during learning and open-book test taking. *British Journal of Educational Psychology, 84*, 329–348. https://doi.org/10.1111/bjep.12021

Adesope, O. O., & Nesbit, J. C. (2012). Verbal redundancy in multimedia learning environments: A meta-analysis. *Journal of Educational Psychology, 104*, 250–263. https://doi.org/10.1037/a0026147

Agarwal, P. K., Nunes, L. D., & Blunt, J. R. (2021). Retrieval practice Consistently benefits student learning: A systematic review of applied research in schools and Classrooms. *Educational Psychology Review, 33*, 1409–1453. https://doi.org/10.1007/s10648-021-09595-9

Baars, M., Wijnia, L., de Bruin, A., & Paas, F. (2020). The relation between students' effort and monitoring judgments during learning: A meta-analysis. *Educational Psychology Review, 32*, 979–1002. https://doi.org/10.1007/s10648-020-09569-3

Bai, C., Yang, J., & Tang, Y. (2022). Embedding self-explanation prompts to support learning via instructional video. *Instructional Science, 50*, 681–701. https://doi.org/10.1007/s11251-022-09587-4

Berthold, K., Nückles, M., & Renkl, A. (2007). Do learning protocols support learning strategies and outcomes? The role of cognitive and metacognitive prompts. *Learning and Instruction, 17*, 564–577. https://doi.org/10.1016/j.learninstruc.2007.09.007

Brame, C. J. (2016). Effective educational videos: Principles and Guidelines for Maximizing student learning from video content. *CBE-Life Sciences Education, 15*. https://doi.org/10.1187/cbe.16-03-0125

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

Cortina, K. S. (2008). Leistungsängstlichkeit. In W. Schneider, & M. Hasselhorn (Eds.), *Handbuch der Psychologie. Handbuch der Pädagogischen Psychologie* (pp. 50–61). Hogrefe.

Dunlosky, J., & Rawson, K. A. (2012). Overconfidence produces underachievement: Inaccurate self evaluations undermine students' learning and retention. *Learning and Instruction, 22*, 271–280. https://doi.org/10.1016/j.learninstruc.2011.08.003

Eitel, A. (2016). How repeated studying and testing affects multimedia learning: Evidence for adaptation to task demands. *Learning and Instruction, 41*, 70–84. https://doi.org/10.1016/j.learninstruc.2015.10.003

Eitel, A., Endres, T., & Renkl, A. (2022). Specific questions during retrieval practice are better for texts containing seductive details. *Applied Cognitive Psychology, 36*, 996–1008. https://doi.org/10.1002/acp.3984

Eitel, A., & Scheiter, K. (2015). Picture or text first? Explaining sequence effects when learning with pictures and text. *Educational Psychology Review, 27*, 153–180. https://doi.org/10.1007/s10648-014-9264-4

Endres, T., Weyreter, S., Renkl, A., & Eitel, A. (2020). When and why does emotional design foster learning? Evidence for situational interest as a mediator of increased persistence. *Journal of Computer Assisted Learning, 36*, 514–525. https://doi.org/10.1111/jcal.12418

Faul, F., Erdfelder, E., Buchner, A., & Lang, A.-G. (2009). Statistical power analyses using G*Power 3.1: Tests for correlation and regression analyses. *Behavior Research Methods, 41*, 1149–1160. https://doi.org/10.3758/BRM.41.4.1149

Fiorella, L., Stull, A. T., Kuhlmann, S., & Mayer, R. E. (2020). Fostering generative learning from video lessons: Benefits of instructor-generated drawings and learner-generated explanations. *Journal of Educational Psychology, 112*, 895–906. https://doi.org/10.1037/edu0000408

Hayes, A. F. (2022). *Introduction to Mediation, Moderation, and Conditional Process Analysis: A Regression-Based Approach* (Third Edition). The Guilford Press.

Hidi, S. (1990). Interest and its contribution as a mental resource for learning. *Review of Educational Research, 60*, 549–571. https://doi.org/10.3102/00346543060004549

Hidi, S. (2001). Interest, reading, and learning: Theoretical and practical considerations. *Educational Psychology Review, 13*, 191–209. https://doi.org/10.1023/A:1016667621114

Hidi, S., & Renninger, K. A. (2006). The four-phase model of interest Development. *Educational Psychologist, 41*, 111–127. https://doi.org/10.1207/s15326985ep4102_4

Hoch, E., Fleig, K., & Scheiter, K. (2023). Can monitoring prompts help to reduce a confidence bias when learning with multimedia? *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie*. https://doi.org/10.1026/0049-8637/a000279

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

Klepsch, M., & Seufert, T. (2021). Making an effort versus experiencing load. *Frontiers in Education, 6*, Article 645284. https://doi.org/10.3389/feduc.2021.645284, 56.

Klingsieck, K. B. (2018). Kurz und knapp – die Kurzskala des Fragebogens „Lernstrategien im Studium" (LIST). *Zeitschrift für Padagogische Psychologie, 32*, 249–259. https://doi.org/10.1024/1010-0652/a000230

Kornell, N., Rhodes, M. G., Castel, A. D., & Tauber, S. K. (2011). The ease-of-processing heuristic and the stability bias: Dissociating memory, memory beliefs, and memory judgments. *Psychological Science, 22*, 787–794. https://doi.org/10.1177/0956797611407929

Krämer, A., & Böhrs, S. (2017). How do Consumers evaluate explainer videos? An empirical study on the effectiveness and Efficiency of different explainer video formats. *Journal of Education and Learning, 6*, 254–266.

Kulgemeyer, C. (2020). A framework of effective science explanation videos informed by criteria for instructional explanations. *Research in Science Education, 50*, 2441–2462. https://doi.org/10.1007/s11165-018-9787-7

Kulgemeyer, C., & Wittwer, J. (2023). Misconceptions in Physics explainer videos and the illusion of understanding: An experimental study. *International Journal of Science and Mathematics Education, 21*, 417–437. https://doi.org/10.1007/s10763-022-10265-7

List, A., & Ballenger, E. E. (2019). Comprehension across mediums: The case of text and video. *Journal of Computing in Higher Education, 31*, 514–535. https://doi.org/10.1007/s12528-018-09204-9

Liu, C., & Elms, P. (2019). Animating student engagement: The impacts of cartoon instructional videos on learning experience. *Research in Learning Technology, 27*. https://doi.org/10.25304/rlt.v27.2124

Lloyd, S. A., & Robertson, C. L. (2012). Screencast tutorials Enhance student learning of statistics. *Teaching of Psychology, 39*, 67–71. https://doi.org/10.1177/0098628311430640

Mason, L., Tarchi, C., Ronconi, A., Manzione, L., Latini, N., & Bråten, I. (2022). Do medium and Context Matter when learning from multiple complementary Digital texts and videos? *Instructional Science, 50*, 653–679. https://doi.org/10.1007/s11251-022-09591-8

Mayer, R. E. (2021). Evidence-based principles for how to design effective instructional videos. *Journal of Applied Research in Memory and Cognition, 10*, 229–240. https://doi.org/10.1016/j.jarmac.2021.03.007

Mayer, R. E. (2022). Cognitive Theory of Multimedia Learning. In R. E. Mayer, & L. Fiorella (Eds.), *Cambridge handbooks in psychology. The Cambridge handbook of multimedia learning* (Third edition, pp. 57–72). Cambridge University Press. https://doi.org/10.1017/9781108894333.008.

Mayer, R. E., & Fiorella, L. (Eds.). (2022). *Cambridge handbooks in psychology. The Cambridge handbook of multimedia learning* (3rd ed.). Cambridge University Press. https://doi.org/10.1017/9781108894333.

Mayer, R. E., Fiorella, L., & Stull, A. (2020). Five ways to increase the effectiveness of instructional video. *Educational Technology Research & Development, 68*, 837–852. https://doi.org/10.1007/s11423-020-09749-6

Moos, D. C., & Bonde, C. (2016). Flipping the classroom: Embedding self-regulated learning prompts in videos. *Technology, Knowledge and Learning, 21*, 225–242. https://doi.org/10.1007/s10758-015-9269-1

Müller, N. M., & Seufert, T. (2018). Effects of self-regulation prompts in hypermedia learning on learning performance and self-efficacy. *Learning and Instruction, 58*, 1–11. https://doi.org/10.1016/j.learninstruc.2018.04.011

Noetel, M., Griffith, S., Delaney, O., Harris, N. R., Sanders, T., Parker, P., del Pozo Cruz, B., & Lonsdale, C. (2022). Multimedia design for learning: An overview of reviews with meta-meta-analysis. *Review of Educational Research, 92*, 413–454. https://doi.org/10.3102/00346543211052329

Noetel, M., Griffith, S., Delaney, O., Sanders, T., Parker, P., Del Pozo Cruz, B., & Lonsdale, C. (2021). Video improves learning in higher education: A systematic review. *Review of Educational Research, 91*, 204–236. https://doi.org/10.3102/0034654321990713

Paik, E. S., & Schraw, G. (2013). Learning with animation and illusions of understanding. *Journal of Educational Psychology, 105*, 278–289. https://doi.org/10.1037/a0030281

Paivio, A. (1986). *Mental representations: A dual coding approach*. Oxford University Press.

Pekrun, R., Vogl, E., Muis, K. R., & Sinatra, G. M. (2017). Measuring emotions during epistemic activities: The epistemically-related emotion scales. *Cognition & Emotion, 31*, 1268–1276. https://doi.org/10.1080/02699931.2016.1204989

Roediger, H. L., & Karpicke, J. D. (2006). Test-enhanced learning: Taking memory tests improves long-term retention. *Psychological Science, 17*, 249–255. https://doi.org/10.1111/j.1467-9280.2006.01693.x

Roelle, J., Schweppe, J., Endres, T., Lachner, A., Aufschnaiter, C. von, Renkl, A., Eitel, A., Leutner, D., Rummer, R., Scheiter, K., & Vorholzer, A. (2022). Combining retrieval practice and generative learning in educational contexts. *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie, 54*, 142–150. https://doi.org/10.1026/0049-8637/a000261

Rost, D. H., Sparfeldt, J. R., & Buch, S. (Eds.). (2018). *Handwörterbuch pädagogische Psychologie (5., überarbeitete und erweiterte Auflage)*. Beltz.

Salomon, G. (1984). Television is "easy" and print is "tough": The differential investment of mental effort in learning as a function of perceptions and attributions. *Journal of Educational Psychology, 76*, 647–658. https://doi.org/10.1037/0022-0663.76.4.647

Schiefele, U. (2009). Situational and individual interest. In K. R. Wentzel, & D. B. Miele (Eds.), *Handbook of motivation at school* (1st ed., pp. 197–222). Routledge.

Schiefele, U., Krapp, A., Wild, K.-P., & Winteler, A. (1993). Der" Fragebogen zum Studieninteresse"(FSI). *Diagnostica, 39*, 335–351.

Schneider, W., & Hasselhorn, M. (Eds.). (2008). *Handbuch der Psychologie. Handbuch der Pädagogischen Psychologie*. Hogrefe.

Schöne, C., Dickhäuser, O., Spinath, B., & Stiensmeier-Pelster, J. (2002). *Skalen zur Erfassung des schulischen Selbstkonzepts : SESSKO*. Hogrefe. https://madoc.bib.uni-mannheim.de/42724.

Senko, C., Perry, A. H., & Greiser, M. (2022). Does triggering learners' interest make them overconfident? *Journal of Educational Psychology, 114*, 482–497. https://doi.org/10.1037/edu0000649

Serra, M. J., & Dunlosky, J. (2010). Metacomprehension judgements reflect the belief that diagrams improve learning from text. *Memory, 18*, 698–711. https://doi.org/10.1080/09658211.2010.506441

Sweller, J., Ayres, P., & Kalyuga, S. (2011). *Cognitive load theory*. Springer.

Szpunar, K. K., Jing, H. G., & Schacter, D. L. (2014). Overcoming overconfidence in learning from video-recorded lectures: Implications of interpolated testing for online education. *Journal of Applied Research in Memory and Cognition, 3*, 161–164. https://doi.org/10.1016/j.jarmac.2014.02.001

Szpunar, K. K., Khan, N. Y., & Schacter, D. L. (2013). Interpolated memory tests reduce mind wandering and improve learning of online lectures. *Proceedings of the National Academy of Sciences of the United States of America, 110*, 6313–6317. https://doi.org/10.1073/pnas.1221764110

Tarchi, C., & Mason, L. (2022). Learning across media in a second language. *European Journal of Psychology of Education, 38*, 1593–1618. https://doi.org/10.1007/s10212-022-00652-7

Tarchi, C., Zaccoletti, S., & Mason, L. (2021). Learning from text, video, or subtitles: A comparative analysis. *Computers & Education, 160*, Article 104034. https://doi.org/10.1016/j.compedu.2020.104034

van der Meij, H., & van der Meij, J. (2014). A comparison of paper-based and video tutorials for software learning. *Computers & Education, 78*, 150–159. https://doi.org/10.1016/j.compedu.2014.06.003

Wang, Y., Wang, F., Mayer, R. E., Hu, X., & Gong, S. (2023). Benefits of prompting students to generate summaries during pauses in segmented multimedia lessons.

Journal of Computer Assisted Learning, 39, 1259–1273. https://doi.org/10.1111/jcal.12797

Wilde, M., Bätz, K., Kovaleva, A., & Urhahne, D. (2009). Überprüfung einer Kurzskala intrinsischer Motivation (KIM). *Zeitschrift für Didaktik der Naturwissenschaften, 15*.

Wittwer, J., & Renkl, A. (2008). Why instructional explanations often do not Work: A framework for understanding the effectiveness of instructional explanations. *Educational Psychologist, 43*(1), 49–64. https://doi.org/10.1080/00461520701756420

Wolf, K. D. (2015). Produzieren Jugendliche und junge Erwachsene ihr eigenes Bildungsfernsehen? Erklärvideos auf Youtube. *Televizion, 28*, 35–39.

Wolf, K. D., Cwielong, I. A., Kommer, S., & Klieme, K. E. (2021). Leistungsoptimierung von Schülerinnen und Schülern durch schulbezogene Erklärvideonutzung auf YouTube: Entschulungsstrategie oder Selbsthilfe? *MedienPädagogik: Zeitschrift Für Theorie Und Praxis Der Medienbildung, 42*, 380–408. https://doi.org/10.21240/mpaed/42/2021.12.31.X

Zheng, L. (2016). The effectiveness of self-regulated learning scaffolds on academic performance in computer-based learning environments: A meta-analysis. *Asia Pacific Education Review, 17*, 187–202. https://doi.org/10.1007/s12564-016-9426-9

Zheng, X., Ma, Y., Yue, T., & Yang, X. (2023). Effects of different types of cues and self-explanation prompts in instructional videos on deep learning: Evidence from multiple data analysis. *Educational Technology Research & Development, 71*, 807–831. https://doi.org/10.1007/s11423-023-10188-2