American Psychological Association

Journal of Educational Psychology

ISSN: 0022-0663

2020, Vol. 112, No. 6, 1273–1283
http://dx.doi.org/10.1037/edu0000414

A Change of Scenery: Does the Setting of an Instructional Video Affect Learning?

Martin Merkt
German Institute for Adult Education and Knowledge Media Research Center

Sabrina Lux
Knowledge Media Research Center

Vincent Hoogerheide
Utrecht University

Tamara van Gog
Utrecht University and University of Tübingen

Stephan Schwan
Knowledge Media Research Center and University of Tübingen

Two experiments investigated the effects of an instructional video's setting on learners' retention and application of the video content. Experiment 1 explored competing hypotheses based on theoretical assumptions about whether an authentic setting would serve as a distraction or as a cue for the instructor's expertise. Participants ($N = 59$) watched a video about floral diagrams and floral formulas that was either shot in a greenhouse (authentic setting) or in front of a white wall (neutral setting). Results showed a beneficial effect of the authentic setting on retention, but not on the application of the video content. Experiment 2 aimed to replicate and extend these findings by investigating whether reinstating the authentic setting as a contextual cue during the test phase would further improve test performance. After watching a video that was either shot in an authentic or a neutral setting, participants ($N = 149$) worked on the retention and application test while a screenshot of the authentic or the neutral setting was presented as a background. Contrary to our expectations, the effect of setting on retention did not replicate and there was no evidence for context effects, despite using the same learning materials and a comparable sample as in Experiment 1. Findings are discussed with regard to potential boundary conditions.

Educational Impact and Implications Statement

Instructional videos are immensely popular, yet design guidelines are scarce. Two experiments examined whether the effectiveness of instructional videos would depend on the setting (i.e., the location) in which instructional videos are recorded. Whereas Experiment 1 showed that students remembered more information about flowers after watching a video recorded in an authentic setting (i.e., greenhouse) than after watching a video recorded in a neutral setting (i.e., office), this setting effect was not replicated in Experiment 2. These findings therefore do not provide convincing evidence that the setting of instructional videos is important enough to be taken into consideration in the design of instructional videos—at least when it is static. Future research has to uncover whether the setting is a more important factor if there is background movement or if there is an actual mismatch between the setting and the learning materials.

Keywords: videos, learning, instructional design, setting, multimedia learning

This article was published Online First October 17, 2019.

Martin Merkt, Department of Research Infrastructures, German Institute for Adult Education, and Department of Realistic Depictions, Knowledge Media Research Center; Sabrina Lux, Department of Realistic Depictions, Knowledge Media Research Center; Vincent Hoogerheide, Department of Educational Sciences, Utrecht University; Tamara van Gog, Department of Educational Sciences, Utrecht University, and LEAD Graduate School and Research Network, University of Tübingen; Stephan Schwan, Department of Realistic Depictions, Knowledge Media Research Center, and Department of Psychology, University of Tübingen.

We thank our video instructor for her efforts during the production of the instructional videos. Further, we appreciate the support of Berit Zimmerling and Danea Keuchel in coding the participants' answers to the open-ended questions. Finally, we thank Selina Säger for coding the participants' answers and drawing the examples of the application task that are depicted in this article.

Correspondence concerning this article should be addressed to Martin Merkt, Department of Research Infrastructures, German Institute for Adult Education, Heinemannstraße 12-14, 53175 Bonn, Germany. E-mail: merkt@die-bonn.de

The popularity of instructional videos has spurred research on how to design videos in order to optimize learning (e.g., Beege, Schneider, Nebel, & Rey, 2017; Boucheix, Gauthier, Fontaine, & Jaffeux, 2018; de Koning, Hoogerheide, & Boucheix, 2018; Fiorella, Van Gog, Hoogerheide, & Mayer, 2017; Hoogerheide, Van Wermeskerken, Loyens, & Van Gog, 2016; Kay, 2012; Merkt, Ballmann, Felfeli, & Schwan, 2018; Merkt, Weigand, Heier, & Schwan, 2011). Even though the design of instructional videos has been on the agenda of educational psychology in recent years, studies have mostly focused on the question of whether students' learning depends on characteristics of the instructor (e.g., Beege et al., 2017; Hoogerheide et al., 2016), the perspective of the camera (e.g., Boucheix et al., 2018; Fiorella et al., 2017), or the presence versus absence of interactive features (e.g., Merkt et al., 2011).

Instructional video research has neglected the potential importance of where the video is recorded. In theater and film theory, the physical location in which a play is situated is referred to as setting (Serlio, 1991). Following this definition, setting is defined as the physical environment in which an instructional video is shot. This definition of setting is also in line with the definition of physical environments as the contextual factors in which learning materials are presented, including ". . . sensory stimuli from the environment that can be perceived by human senses . . ." (Choi, Van Merriënboer, & Paas, 2014, p. 229). This definition entails auditory and visual information including ". . . physical properties such as the background color of a computer screen . . ." (Choi et al., 2014, p. 230). Thus, the setting of an instructional video must be considered a pervasive design aspect that is a part of the physical environment of the instructor as well as the learner, yet not a part of the learning task itself. To clarify, for example, in a video tutorial about plants, a greenhouse setting constitutes a part of the physical learning environment, whereas plants that are used to demonstrate specific aspects of the learning materials are part of the learning task. The present study addresses the effect of setting by investigating whether shooting an instructional video in an authentic setting affects how well learners remember and apply the video's contents.

Whether shooting instructional videos in an authentic setting would benefit or hamper learning is an open question. Three competing hypotheses can be identified about how the setting of an instructional video may affect learning. First, an authentic setting that fits the contents of learning materials may serve as a persistent cue of an instructor's expertise, comparable to an instructor's age (Hoogerheide et al., 2016) or clothing (Glick, Larsen, Johnson, & Branstiter, 2005; Morris, Gorham, Cohen, & Huffman, 1996). Second, an authentic setting may serve as a distraction that reduces learners' attention to the relevant contents (i.e., seductive details effect; Harp & Mayer, 1998; Rey, 2014). In addition, research on context congruency effects (Huff, Maurer, & Merkt, 2018; Smith & Sinha, 1987; Smith & Vela, 2001) implies that the effect of instructional video's setting may be qualified by the context that is provided during the test phase. In the following sections, we elaborate on these different lines of research that come to competing predictions on the effect of an instructional video's setting on learning.

Setting as a Cue for Expertise

The context in which humans encounter people or objects shapes their judgments about these people or objects (Schwarz, 2007). In particular, context may prime concepts that are in turn attributed to people (Adam & Galinsky, 2012). Clothing was investigated as such a contextual factor that influences judgments about people's expertise. For instance, the same teachers were rated to be more intelligent when they wore formal clothes than when they wore casual clothes (Morris et al., 1996). Given that contextual factors may affect judgments about people, it is likely that the physical environment (i.e., setting) in which a video was produced would have similar effects. In particular, a physical environment that fits the learning materials should be considered an authentic place in which experts would usually work.

If an authentic setting actually works as a cue for an instructor's expertise, this perceived expertise may affect how much students learn. There are some indications from eye tracking research that learners pay more attention to experts than to novices (see, e.g., Cheng, Tracy, Foulsham, Kingstone, & Henrich, 2013). Further, students learned more from experts than from (advanced) students (Boekhout, van Gog, van de Wiel, Gerards-Last, & Geraets, 2010; Lachner & Nückles, 2015); however, most of these studies also varied the contents of the instructors' explanation as part of their experimental manipulation (Hoogerheide, Loyens, Jadi, Vrins, & Van Gog, 2017). Hoogerheide et al. (2016) did control for the content of the instructional video. They manipulated whether the instructor was a peer or an adult and introduced herself as someone with low versus high expertise. There was no effect of purported expertise on learning outcomes. However, participants rated the purported experts' explanations as being of higher quality than nonexperts' explanations. Note that whereas one could argue that these findings contradict the idea that purported expertise might matter for learning, the expertise manipulation was not salient throughout the entire video, but only during a short introduction. Moreover, the videos including an adult instructor were perceived to provide a better explanation and resulted in better learning outcomes than videos with a peer instructor. Hoogerheide and colleagues suggested that age might have served as a cue for the instructor's expertise that was salient throughout the entire video because the perceived fit between the instructors' age and the learning domain (i.e., troubleshooting electrical circuits) might have been higher for adult than for adolescent instructors. It is plausible that an authentic setting would also serve as a persistent cue for expertise, comparable to an instructor's age.

Setting as a Distraction

Cognitive theory of multimedia learning (Mayer, 2014) and Cognitive Load Theory (Sweller, Van Merriënboer, & Paas, 1998) argue that learning materials should be designed to facilitate processing of relevant information in order to make optimal use of learners' limited cognitive resources (Mayer, 2014; Sweller et al., 1998). Whereas it is important to adhere to different design principles that are suggested in the context of these theories, this section focuses on the seductive detail effect postulating that learning materials should not be enriched with irrelevant details that may distract learners from relevant contents of the learning materials (Harp & Mayer, 1998; Rey, 2012). Following this rationale, an authentic setting might hamper students' learning because the setting of an instructional video is not part of the learning task. Remember that in line with Choi et al. (2014), the setting of a video was defined as the physical environment of an instructor that is not part of the actual learning task. In particular, when shooting on location in an authentic setting, the setting may include distinct visual details that capture learners' attention. Thus, presenting an authentic setting in an instructional video might constitute a seductive detail that piques the learners' interest but also distracts their attention (see Garner, Brown, Sanders, & Menke, 1992; Harp & Mayer, 1998).

Even though decorative pictures (as one example of seductive details), compared to instructional pictures, may have a positive effect on learners' affective states (Lenzner, Schnotz, & Müller, 2013; for a comprehensive overview of different types and functions of pictures, see Carney & Levin, 2002), a meta-analysis by Rey (2012) paints a less favorable picture about adding decorative elements to learning materials. In particular, Rey (2012) identified small to medium negative effects of seductive details on learning outcomes such as retention or transfer. These negative effects of seductive details were often attributed to distracting learners' attention from relevant pieces of information in the learning materials (Lehman, Schraw, McCrudden, & Hartley, 2007; Sanchez & Wiley, 2006) or to activating inappropriate schemata that guide learners' representation of learning materials (Harp & Mayer, 1998). In support of the assumption that seductive details hamper learning because they attract learners' attention and thus distract attention from relevant contents, seductive details were more harmful to learners with low attention control (Rey, 2014; Sanchez & Wiley, 2006).

According to Mayer (2014), distraction in transient media (e.g., instructional videos) hampers learning because relevant information may no longer be available for subsequent stages of information processing if the information was not attended to at exactly the moment at which the relevant piece of information was presented. These subsequent stages of information processing entail the mental organization and integration of verbal and visual information as well as the integration of newly learned information with prior knowledge in order to create a comprehensive mental representation of the learning materials (Mayer, 2014). Further, information that is irrelevant for learning outcomes (e.g., setting) may hamper learning because it draws on limited working memory resources so that capacity for essential processing of relevant materials is reduced.

Setting as a Contextual Retrieval Cue

The setting of an instructional video can be considered context that is encoded together with the actual learning materials. As such, the setting of an instructional video may also be considered a contextual retrieval cue if the same context is successfully reinstated during the test phase. There is some evidence for such context congruency effects in the research literature (Huff et al., 2018; Smith & Sinha, 1987; Smith & Vela, 2001). The explanation of these effects usually relies on a shared assumption postulated in multiple theories about human memory (e.g., encoding specificity assumption; Tulving & Thomson, 1973), stating that retrieval improves with an increased overlap of characteristics of the learning phase and test phase (also see Smith & Vela, 2001). In particular, the encoding specificity assumption states that retrieval should be improved if contextual factors during the learning phase are similar to contextual factors during the test phase. Similarly, the concept of transfer-appropriate processing would imply that performance on knowledge tests would improve if processes that are relevant for information retrieval are already part of the learning phase (Morris, Bransford, & Franks, 1977). This effect should be more pronounced for feature-rich settings than for settings of lower visual complexity because context reinstatements were most likely to improve retention if learners generate unique associations between the learning materials and contextual information during the learning phase (Doss, Picart, & Gallo, 2018; Murnane, Phelps, & Malmberg, 1999).

One of the best-known demonstrations of context effects was provided by Godden and Baddeley (1975). They showed that memory for word lists was better when divers who learned word lists under water were tested under water, whereas people who learned word lists on land showed better memory when they were tested on land. Similar context effects were observed for background music (Balch, Bowman, & Mohler, 1992), smells (Cann & Ross, 1989), bodily postures (Rand & Wapner, 1967), and engaging in motor activities (Huff et al., 2018). However, the most convincing evidence in favor of the assumption that an instructional video's setting would serve as a contextual retrieval cue was provided by Smith and colleagues (Smith, Handy, Angello, & Manzano, 2014; Smith & Manzano, 2010). Smith and Manzano (2010) had participants learn word lists that were displayed word by word in the center of the screen and superimposed on unrelated video clips that were shown in the background of the computer screen. Participants remembered the word lists better when a congruent background video was presented during the learning phase and test phase. Thus, a video background may serve as a contextual cue that facilitates retrieval of information during the test phase.

Overview of the Hypotheses

The literature reviewed in the previous sections allows for the derivation of three partly competing hypotheses that are specified in the following paragraphs. For easier referencing throughout the remaining article, these hypotheses are labeled as *expertise hypothesis*, *distraction hypothesis*, and *retrieval cue hypothesis*.

Expertise Hypothesis

In line with the assumption that an authentic setting in an instructional video serves as a permanent cue for the instructor's expertise, an authentic setting should result in superior learning outcomes than a neutral setting.

Distraction Hypothesis

Based on the assumption that an authentic setting may serve as a distraction that hinders the processing of relevant information, a neutral setting should result in better learning outcomes than an authentic setting.

Retrieval Cue Hypothesis

In line with the assumption that an instructional video's setting may serve as contextual information that is encoded together with the learning materials, there should be larger effects of an authentic setting if contextual information is reinstated as a contextual retrieval cue in the test phase.

Two experiments tested these hypotheses. Experiment 1 was designed to test the *expertise hypothesis* against the *distraction hypothesis*. Experiment 2 was designed to replicate the findings of Experiment 1 and to additionally test the *retrieval cue hypothesis*.

Experiment 1

To explore the effect of an instructional video's setting on learning outcomes, participants watched an instructional video about flowers, floral diagrams, and floral formulas that was shot in a greenhouse in a botanical garden (authentic setting) or in front of a white wall in an office (neutral setting). The context during the test phase was kept constant (i.e., neutral) across all conditions. The experiment served to test the *expertise hypothesis* that favors the authentic setting against the *distraction hypothesis* that favors the neutral setting.

Method

Participants and design. Overall, 65 participants from a database of a German research institute took part in this experiment. However, six participants were excluded from the analyses: four because their scores were lower than 10% on one or both of the knowledge tests (suggesting they did not seriously study the materials), one because (s)he aborted the experiment, and one because (s)he studied biology. Hence, the final sample consisted of 59 participants (40 female, 19 male). Fifty-five participants were enrolled in different majors at a German university, whereas four participants reported having another occupation. Their mean age was 24.05 years ($SD = 3.19$). They were randomly assigned either to watch an instructional video shot in an authentic setting ($n = 29$) or to watch an instructional video shot in a neutral setting ($n = 30$), resulting in a one-factorial between subject-design with two levels.

Video tutorials. Two instructional videos were produced that were equal in content and included the same instructor (see Figure 1). The pace of the narration and the instructor's gestures were parallelized so that the videos did not differ in their coverage of different aspects of the learning materials. In particular, after first producing the video that was shot in the authentic setting, the video shot in the neutral setting was produced with Sabrina Lux paying close attention to the overlap of the two videos. For this purpose, each video was split into different segments. Individual segments were redone until Sabrina Lux was satisfied with the overlap between the two videos, including the use of gestures. Further, in post-production, the length of the individual segments was adjusted so that each individual topic was included for the same duration in both videos.

Both videos included a flipchart (left-hand side from the observer perspective) and a table with plants (right-hand side). Both the flipchart and the plants on the table served as demonstration material and were thus considered part of the learning materials and not part of the setting. The flipchart and table were arranged in a similar way in both videos; however, due to spatial constraints, the angle of the table was slightly different between the videos (see Figure 1). Because there were zoom-ins both on the graphics, flipchart, and plants when relevant aspects of the learning materials were explained, these slight deviations are unlikely to affect learning outcomes and are thus negligible.

The contents of the video tutorials were based on a textbook about plant classification (Lüder, 2004). The instructor described the composition of flowers and how floral diagrams and floral formulas are created using schematic representations on the flipchart. Zoom-ins focused on information on the flipchart whenever the information was relevant for comprehension. Floral diagrams are schematic depictions of flowers in which the structure of the different elements of the flowers as well as their characteristics are shown (Lüder, 2004; Ronse de Craene, 2010). Floral formulas describe the flowers' structure in a mathematical way (i.e., using letters and numbers that represent the elements of the flowers; Lüder, 2004; Prenner, Bateman, & Rudall, 2010). During this part of the video, the instructor used plants that were placed on the table to demonstrate some of the characteristics. Zoom-ins on the plants were used so that individual parts of the plants were clearly visible. The video ended with two examples in which floral diagrams and floral formulas for two different plants were explained in detail. Again, zoom-ins were used to focus learners' attention on relevant information on the flipchart. Each video lasted 13 min and 57 s.

Because the learning materials were concerned with floral diagrams and floral formulas, a greenhouse in a botanical garden was used as the location for the authentic setting. In the background, there were plants in front of which the female instructor, flipchart, and table with plants were standing. Due to a small trickle running

Figure 1. Screenshots from the instructional videos shot in an authentic setting (left) and a neutral setting (right). The instructor's face was not pixelated in the experiments. See the online article for the color version of this figure.

through the greenhouse, there was an audible tinkling of water in the background that was toned down in post-production.

As the neutral setting, an office at a research institute was used. In the background, there was a white wall in front of which the instructor, flipchart, and table with the plants were arranged comparable to the arrangement in the authentic setting. Due to the architecture of the room, there was an audible reverb that was toned down in post-production.

Measures.

Knowledge tests. Learners' retention of the content of the instructional video as well as their ability to apply the acquired knowledge were assessed. The retention test consisted of 28 open-ended questions that required short answers (e.g., "Which part of the flower is considered to be the male part?" Answer: "Stamen."). Based on a coding scheme, one point was awarded for each correct answer, with partial credit (i.e., 0.5) for partially correct answers. Thus, the maximum score for retention was 28. Two independent raters coded all of the participants' answers to all of the questions. Both raters were blind to condition and showed an excellent intraclass correlation (ICC) for the overall scores of each participant, $r = .99$, $p < .001$. Disagreements were resolved by discussion, and the raters' agreed-upon coding was used as the score for retention.

With regard to the application test, participants were required to generate eight floral formulas based on eight floral diagrams and on the corresponding schematic profile of the flowers that were taken from a textbook about plant classification (Lüder, 2004). Please refer to Figure 2 for an illustrative drawing of an item of an application task. Ten points could be earned per formula; the maximum score for the application test was 80. Based on a coding scheme, 0.5 points or 1 point were subtracted from this score for each predefined error that participants made. The formulas generated by all participants were coded by two independent raters who were blind to condition. The raters showed an excellent ICC for the overall scores of each participant, $r = .99$, $p < .001$. Disagreements were resolved by discussion and the agreed-upon coding was used as the score for application.

Subjective rating scales. To shed light on mechanisms that could potentially explain the hypothesized (positive or negative) relationship between setting and learning, participants were asked various questions after the learning phase. More specifically, participants were asked to provide their view on a scale of 1 (*not at all/very low*) to 7 (*a lot*) concerning: distraction ("How distracting was the setting of the video?"), difficulty ("How difficult were the contents of the video?"), comprehension ("How much did the tutorial help you to understand the structure of floral diagrams and floral formulas?"), quality of the instruction ("How would you rate the quality of the explanation with regard to the learning contents?"), professionalism of the video ("How professional did the video appear to be?"), expertise of the instructor ("How much of an expert was the instructor?"), joy of learning ("How much did you enjoy learning with an instructional video (instead of a text)?"), motivation ("Would you like to see similar tutorials about other topics?"), and interestingness ("How interesting did you find the contents of the tutorial?"). Further, participants' self-efficacy was assessed with two items on a scale of 1 (*not at all*) to 7 (*a lot*). The items were "After seeing the tutorial, how capable did you feel about reading floral diagrams?" and "After seeing the tutorial, how capable did you feel about creating floral formulas?" Cronbach's alpha for the two items was .88; therefore, the two ratings were averaged to one composite score for self-efficacy. Finally, cognitive load was measured on a scale ranging from 1 (*not effortful at all*) to 9 (*very effortful*) directly after the video, after the retention test, and after the application test ("How much effort did you invest in following the contents of the tutorial/answering the questions/creating the floral formulas?"; see Paas, 1992).

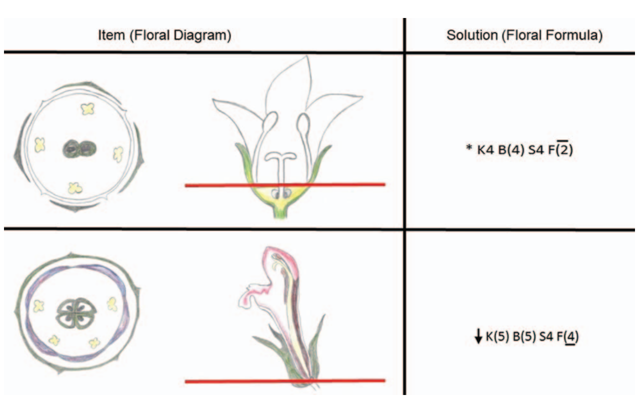

Figure 2. Illustrative drawing of two items of the application task (left) and the solutions (right). See the online article for the color version of this figure.

Other variables and manipulation check. To check whether the experimental groups were comparable, we assessed participants' interest ("How interested are you in the topic of plant classification?") and subjective prior knowledge ("How much prior knowledge do you have about plant classification?") before the learning phase on scales ranging from 1 (*not at all/no prior knowledge*) to 7 (*a lot/a lot of prior knowledge*). Furthermore, participants were asked to give a definition of floral diagrams and floral formulas in two open-ended questions that were analyzed according to a predefined coding scheme. The scores for the two open-ended questions were added for an objective prior knowledge score (maximum score = 16).

As a manipulation check, participants were asked to indicate 'How would you rate the thematic fit between the video's content and the setting of the tutorial?' on a scale from 1 (*very low*) to 7 (*very high*). Further, participants rated the perceived quality of the instructional video by answering the question "How would you rate the quality of the video material" on a scale from 1 (*very low quality*) to 7 (very high quality).

Procedure. The experiment took place in a research laboratory. Up to four participants took part in the experiment simultaneously, but all participants worked on individual computers at their own pace. After prior knowledge (subjective and objective) and prior interest were assessed, participants watched one of the two instructional videos, depending on the experimental condition to which they were randomly assigned. The videos were presented with a size of 768 × 576 pixels on a 27-in. screen with a screen resolution of 1920 × 1080 pixels. Audio was presented via headphones. In the computer-based instructions, participants were told to watch the video attentively and to try to comprehend the contents. Moreover, they were asked to remember as much information as possible for a subsequent knowledge test. After participants watched the video, they were asked to rate cognitive load and to give their subjective ratings about the video, the video quality, and the thematic fit between the video's setting and the learning materials. Then, participants worked on the retention test, followed by the retrospective assessment of cognitive load during the retention test. Afterward, they worked on the application test, followed by another rating of cognitive load. Finally, demographic data were collected. During the entire post questionnaire, the questions (including the knowledge test) were presented on a default light blue background of the experimental software MediaLab, 2012 (Jarvis, 2012). Overall, the study took about an hour. Participants received financial compensation for their efforts. This procedure was approved by the local ethics committee.

Results

Data were analyzed with ANOVAs with the independent variable setting (authentic vs. neutral) and the respective outcomes as dependent variables. The significance level was set at $\alpha = .050$. Descriptive data for all analyses are provided in Table 1.

Manipulation check and individual differences. As expected, the manipulation check revealed that the authentic setting was rated to be a better fit for the topic of the learning materials ($M = 5.69$, $SD = 1.00$) than the neutral setting ($M = 4.83$, $SD = 0.95$), $F(1, 57) = 11.34$, $p = .001$, $\eta_p^2 = .17$. Judgments of the video quality did not differ significantly between the authentic ($M = 3.34$, $SD = 1.17$) and neutral setting ($M = 3.87$, $SD = 1.28$), $F(1, 57) = 2.66$, $p = .108$, $\eta_p^2 = .05$. There were no differences between groups on learners' self-reported interest in and prior knowledge (both subjective and objective) about plant classification, all Fs < 1.

Knowledge tests. With regard to the main dependent variables, there was an effect of setting on retention, $F(1, 57) = 6.42$, $p = .014$, $\eta_p^2 = .10$. In particular, the experimental group that studied the video recorded in an authentic setting ($M = 18.12$, $SD = 5.27$, $\%_{correct} = 64.71$) outperformed the control group that studied the video shot in a neutral setting ($M = 14.73$, $SD = 5.00$, $\%_{correct} = 52.61$).

In contrast, there was no significant effect of setting on knowledge application, $F(1, 57) = 1.99$, $p = .164$, $\eta_p^2 = .03$ (authentic setting: $M = 67.93$, $SD = 13.20$, $\%_{correct} = 84.91$; neutral setting: $M = 62.80$, $SD = 14.70$, $\%_{correct} = 78.50$).

Subjective rating scales. The setting of the instructional video did not affect perceived difficulty and perceived cognitive load measured directly after watching the video, directly after the retention test, and directly after the application test, all Fs < 1. Moreover, participants in the authentic setting group ($M = 3.07$, $SD = 1.75$) did not feel more distracted by the video's setting than participants in the neutral setting group ($M = 2.40$, $SD = 1.43$), $F(1, 57) = 2.59$, $p = .113$, $\eta_p^2 = .04$. Further, there was no effect with regard to perceived comprehension, $F(1, 57) = 3.05$, $p = .086$, $\eta_p^2 = .05$, with learners that watched the video shot in an authentic setting ($M = 5.79$, $SD = 1.01$) reporting similar levels of comprehension as the learners that had watched the video shot in a neutral setting ($M = 5.30$, $SD = 1.15$).

With regard to the perceived expertise of the instructor, the quality of the explanation, and the professionalism of the tutorial, there were no significant differences between the two groups, all Fs < 1. Finally, the two groups did not differ with regard to self-reported learning enjoyment, $F < 1$, motivation, $F < 1$, interest, $F(1, 57) = 1.43$, $p = .236$, $\eta_p^2 = .03$, and self-efficacy, $F(1, 57) = 1.24$, $p = .269$, $\eta_p^2 = .02$.

Table 1
Means (and Standard Deviations in Parentheses) for Experiment 1

Variable name	Authentic setting	Neutral setting	Overall
Fit	5.69 (1.00)	4.83 (.95)	5.25 (1.06)
Video quality	3.34 (1.17)	3.87 (1.28)	3.61 (1.25)
Prior interest	2.59 (1.27)	2.47 (1.38)	2.53 (1.32)
Prior knowledge (subjective)	2.21 (1.26)	2.17 (0.91)	2.19 (1.09)
Prior knowledge (objective)	0.16 (0.48)	0.15 (0.53)	0.15 (0.50)
Retention	18.12 (5.27)	14.73 (5.00)	16.40 (5.37)
Application	67.93 (13.20)	62.80 (14.70)	65.32 (14.10)
Difficulty	3.38 (1.32)	3.57 (1.17)	3.47 (1.24)
Mental effort: Video	4.72 (2.30)	5.00 (1.98)	4.86 (2.13)
Mental effort: Retention	4.17 (2.14)	4.10 (1.86)	4.14 (1.99)
Mental effort: Application	4.66 (2.09)	4.83 (2.12)	4.75 (2.09)
Distraction	3.07 (1.75)	2.40 (1.43)	2.73 (1.62)
Comprehension	5.79 (1.01)	5.30 (1.15)	5.54 (1.10)
Instructor's expertise	4.97 (0.91)	5.00 (1.26)	4.98 (1.09)
Quality of the explanation	5.28 (0.96)	5.27 (1.08)	5.27 (1.01)
Professionalism	3.62 (1.45)	3.90 (1.35)	3.76 (1.40)
Joy of learning	5.00 (1.39)	5.00 (1.11)	5.00 (1.25)
Motivation	4.59 (1.50)	4.53 (1.80)	4.56 (1.64)
Interestingness	4.52 (1.53)	4.07 (1.36)	4.29 (1.45)
Self-efficacy	4.59 (1.05)	4.25 (1.25)	4.42 (1.16)

Discussion

The main aim of Experiment 1 was to contrast the *expertise hypothesis*, which argues that an authentic setting should facilitate learning because it is a cue for an instructor's expertise, with the *distraction hypothesis*, which states that an authentic setting should hamper learning because it distracts the learners' attention from the relevant contents of the learning materials. Results provided only partial support for the pattern predicted by the *expertise hypothesis*. In particular, in line with this hypothesis, students who viewed an instructional video recorded in an authentic setting performed significantly higher on a retention test (65% correct) than those who watched the neutral setting video (53% correct). Descriptively, participants who watched the video shot in an authentic setting (85% correct) performed better on the application task than participants who watched the video shot in a neutral setting (79% correct), but this difference was not statistically significant. However, because the setting of the video did not affect students' perceptions of the instructor's expertise, the experiment does not support the assumption that this effect on retention is driven by how students perceive the expertise of the instructor. As for the distraction hypothesis, our findings suggest that this does not apply: An authentic setting did not hamper learning outcomes, and learners' subjective ratings of distraction did not differ between conditions.

Results provided only partial evidence for the effect of an instructional video's setting on retention. Further, Experiment 1 can be considered exploratory for testing two competing hypotheses regarding learning outcomes. Therefore, it was decided to replicate these findings in Experiment 2 with a larger sample. The decision to increase the sample size was based on a post hoc power analysis with GPower 3.1.9.2. (Faul, Erdfelder, Buchner, & Lang, 2009) was 71%, whereas the achieved power to detect small to medium effects ($\eta_p^2 = .06$) was 48%.

Further, Experiment 2 not only aimed to replicate but also to expand on Experiment 1 by testing whether the beneficial effect of an authentic setting could be increased by providing learners with an image of the same authentic setting as a retrieval cue during the test phase. In particular, the setting of an instructional video can be considered context that is present while learners are processing learning materials and thus encoded together with learning materials. As such, setting may constitute a contextual retrieval cue that facilitates performance on knowledge tests if the context is reinstated while testing (Huff et al., 2018; Smith & Sinha, 1987; Smith & Vela, 2001). Therefore, in Experiment 2, the setting of the instructional video was reinstated by including a picture that was shot in the same location as the background picture in the test phase. If learners formed unique associations between the setting of the instructional video and the learning materials, reinstating contextual information should result in larger effects of an authentic setting on retention and application of learning materials.

Experiment 2

Experiment 2 investigated whether the findings of Experiment 1 would replicate. Moreover, Experiment 2 tested whether the setting of an instructional video would improve learning by serving as a contextual retrieval cue during the test phase.

We expected to replicate the findings of Experiment 1. In particular, participants watching the video shot in an authentic setting should outperform participants watching the video shot in a neutral setting on retention (but not necessarily on application). Further, in line with the *retrieval cue hypothesis* stating that an instructional video's setting may serve as contextual information that is encoded together with the learning materials, larger effects of setting should be observed in the condition in which the authentic setting is activated as a contextual retrieval cue in the test phase.

Method

Participants and design. A power analysis with GPower 3.1.9.2. (Faul et al., 2009) was conducted with the goal to reliably detect the effect of setting on retention even with a small to medium effect size ($\eta_p^2 = .06$). This analysis revealed that 125 participants were necessary to replicate the effect of setting during the learning phase on retention, with α set to .05 and 80% power. Data from 160 participants were collected to account for potential drop-outs and the possibility that the effect would be smaller. For sampling, the same database as for the first experiment was used; however, participants who had already taken part in Experiment 1 were not contacted. Seven participants had to be excluded from analyses because their scores were lower than 10% on one or both of the two knowledge tests, one participant because (s)he indicated having finished an apprenticeship in gardening, one because (s)he indicated knowing the instructor, and two because they studied biology. The remaining 149 participants' (115 female, 34 male) mean age was 23.20 ($SD = 3.50$). One-hundred-and-forty-five participants were enrolled in different majors at a German university, whereas four participants indicated another occupation. Participants were randomly assigned to one of the four conditions that resulted from the between-subjects variables 'setting during the learning phase' (authentic vs. neutral) and 'setting during the test phase' (authentic vs. neutral).

Materials. The video tutorials were the same as in Experiment 1. However, this experiment differed from Experiment 1 in that pictures of the settings of the instructional videos were included as background during the knowledge tests in Experiment 2 in order to reinstate the context of the learning phase (see Figure 3). Background pictures were taken in the same locations as the videos but did not include the learning materials (i.e., flipchart, table with plants).

Measures. The measures were the same as in Experiment 1. Regarding the knowledge tests, two independent raters coded the answers of 60 participants for the retention ($ICC > .99$, $p < .001$) and the application task ($ICC = .97$, $p < .001$). Because correlations were very high, one rater finished coding the remaining tests. Scores of the rater coding all participants were used in the reported analyses. With regard to participants' subjective ratings, Cronbach's alpha for the two self-efficacy items was .88; therefore, the two items were combined into one score for Experiment 2 as well.

Procedure. The procedure for Experiment 2 was comparable to Experiment 1 with two slight deviations. First, the experiment was run with up to five participants on individual computers concurrently, compared to four concurrent participants in Experiment 1. Second, a picture of the authentic setting or the neutral

Figure 3. Exemplary item of the retention test that was presented against the background of the authentic setting (left) or the neutral setting (right). See the online article for the color version of this figure.

setting was presented as a background during the knowledge tests. Because participants' subjective ratings of the video were collected before the two knowledge tests, these ratings were not affected by the second independent variable (i.e., setting during the test phase), as this second factor had not yet been manipulated when participants gave these ratings.

Results

For the main analyses regarding participants' performance on the knowledge tests (including mental effort ratings) as well as for prior interest and prior knowledge, 2 × 2 ANOVAs were conducted with two between subject variables, namely setting during the learning phase (authentic vs. neutral) and setting during the test phase (authentic vs. neutral). For learners' subjective ratings of the instructional videos and for the manipulation check, one-factorial ANOVAs were conducted with the factor setting during the learning phase because participants had not started on the posttests yet when the subjective ratings were assessed. The significance level was set to $\alpha = .050$. Descriptive data for all variables are presented in Table 2.

Manipulation check and individual differences. The manipulation check revealed the expected main effect of setting during the learning phase, $F(1, 147) = 28.91$, $p < .001$, $\eta_p^2 = .16$, with participants rating the authentic setting ($M = 5.78$, $SD = 1.11$) to be more fitting for the learning materials than the neutral setting ($M = 4.64$, $SD = 1.46$). With regard to the quality of the video materials, there also was a main effect of setting during the learning phase, $F(1, 147) = 4.45$, $p = .037$, $\eta_p^2 = .03$; however, with a small effect size. Participants who watched the video shot in the neutral setting ($M = 3.48$, $SD = 1.34$) judged it to be of a higher quality than those participants who watched the video shot in the authentic setting ($M = 3.04$, $SD = 1.20$).

With regard to self-reported prior interest, there were no main effects of setting during the learning phase, $F(1, 145) = 1.13$, $p = .290$, $\eta_p^2 = .01$, setting during the test phase, $F < 1$, and no interaction between the two variables, $F(1, 145) = 1.03$, $p = .313$, $\eta_p^2 = .01$. Regarding the learners' prior knowledge (subjective and objective), there also were no main effects and no interaction, all Fs < 1.

Knowledge tests. The main analyses were concerned with learners' performance on the knowledge tests. Contrary to expectations, with regard to retention ($M = 17.26$, $SD = 5.54$, $\%_{correct} = 61.64$), there were no main effects of setting during the learning phase, $F(1, 145) = 1.11$, $p = .293$, $\eta_p^2 = .01$, setting during the test phase, $F < 1$, and no interaction between the two factors, $F < 1$. For learners' self-reported cognitive load after the retention test, there were no main effects of setting during the learning phase and setting during the test phase as well as no interaction between the two factors, all Fs < 1.

For the application of the learned contents ($M = 66.31$, $SD = 13.03$, $\%_{correct} = 82.89$), there also were no main effects of setting during the learning phase, setting during the test phase, and no interaction between the two factors, all Fs < 1. For self-reported cognitive load after the application test, there were no main effects of setting during the learning phase, $F < 1$, and setting during the test phase, $F(1, 145) = 1.70$, $p = .194$, $\eta_p^2 = .01$, nor an interaction between the two factors, $F(1, 145) = 1.73$, $p = .190$, $\eta_p^2 = .01$.

Table 2
Means (and Standard Deviations in Parentheses) for Experiment 2

Learning phase	Authentic setting			Neutral setting			Overall		
Variable name / Test phase	Authentic setting	Neutral setting	Overall	Authentic setting	Neutral setting	Overall	Authentic setting	Neutral setting	Overall
Fit			5.78 (1.11)			4.64 (1.46)			5.21 (1.42)
Video quality			3.04 (1.20)			3.48 (1.34)			3.26 (1.29)
Prior interest	2.81 (1.56)	2.59 (1.55)	2.70 (1.55)	2.33 (1.24)	2.58 (1.23)	2.45 (1.23)	2.57 (1.42)	2.59 (1.39)	2.58 (1.40)
Prior knowledge (subjective)	2.22 (1.25)	2.30 (1.20)	2.26 (1.22)	2.00 (1.17)	2.17 (1.06)	2.08 (1.11)	2.11 (1.21)	2.23 (1.12)	2.17 (1.17)
Prior knowledge (objective)	0.20 (0.52)	0.41 (1.64)	0.30 (1.21)	0.18 (0.56)	0.29 (0.74)	0.23 (0.65)	0.19 (0.54)	0.35 (1.27)	0.27 (0.97)
Retention	16.81 (6.34)	16.72 (6.18)	16.76 (6.22)	18.51 (4.77)	16.93 (4.69)	17.75 (4.77)	17.68 (5.62)	16.82 (5.46)	17.26 (5.54)
Application	64.70 (14.56)	66.22 (14.43)	65.46 (14.42)	67.60 (11.47)	66.67 (11.78)	67.15 (11.55)	66.19 (13.06)	66.44 (13.10)	66.31 (13.03)
Difficulty			3.54 (1.47)			3.21 (1.35)			3.38 (1.42)
Mental effort: Video			5.20 (2.04)			5.01 (2.06)			5.11 (2.05)
Mental effort: Retention	4.70 (2.48)	4.27 (2.18)	4.49 (2.33)	4.46 (2.33)	4.22 (2.21)	4.35 (2.26)	4.58 (2.39)	4.25 (2.18)	4.42 (2.29)
Mental effort: Application	5.51 (2.39)	4.54 (2.26)	5.03 (2.36)	5.38 (2.28)	5.39 (2.13)	5.39 (2.19)	5.45 (2.32)	4.96 (2.22)	5.21 (2.28)
Distraction			3.01 (1.78)			3.04 (1.77)			3.03 (1.77)
Comprehension			5.41 (1.11)			5.52 (1.33)			5.46 (1.22)
Instructor's expertise			5.11 (1.25)			4.83 (1.35)			4.97 (1.30)
Quality of the explanation			4.95 (1.27)			5.05 (1.30)			5.00 (1.28)
Professionalism			3.47 (1.52)			3.60 (1.58)			3.54 (1.54)
Joy of learning			4.93 (1.47)			4.81 (1.57)			4.87 (1.51)
Motivation			4.47 (1.57)			4.15 (1.80)			4.31 (1.69)
Interestingness			4.00 (1.52)			4.04 (1.61)			4.02 (1.56)
Self-efficacy			4.56 (1.34)			4.47 (1.43)			4.51 (1.39)
N	37	37	74	39	36	75	76	73	149

Subjective rating scales. There was no effect of setting during the learning phase on any of the self-reported measures, namely: difficulty of the video tutorial, $F(1, 147) = 2.00$, $p = .159$, $\eta_p^2 = .01$, cognitive load directly after watching the video, $F < 1$, distraction, $F < 1$, and self-reported comprehension, $F < 1$, expertise of the instructor, $F(1, 147) = 1.75$, $p = .188$, $\eta_p^2 = .01$, quality of the explanation, $F < 1$, professionalism of the video, $F < 1$, learning enjoyment, $F < 1$, motivation, $F(1, 147) = 1.39$, $p = .241$, $\eta_p^2 = .01$, interestingness, $F < 1$, and self-efficacy, $F < 1$.

Discussion

Experiment 2 investigated whether the results of Experiment 1 would replicate and whether the hypothesized beneficial effect of an authentic setting compared to a neutral setting on retention would be enhanced by providing the same authentic setting as a retrieval cue during the test phase. However, the positive effect of an authentic setting on retention found in Experiment 1 did not replicate. Moreover, because there were no differences between groups on any of the outcome variables, the findings of Experiment 2 do not provide any support for the *expertise hypothesis*, the *distraction hypothesis*, or the *retrieval cue hypothesis*. It is unlikely that the student population or the materials could explain why the effect of setting on retention found in Experiment 1 did not replicate in Experiment 2 because the samples of the two experiments were recruited from the same pool of participants and because there were no changes in the learning materials, knowledge tests, and procedure between the two experiments.

Finally, participants who watched the video that was shot in a neutral setting rated the video as being of higher quality than those who watched the video shot in an authentic setting; however, with a small effect size. In this regard, it is important to note that participants only rated the video that they saw in the learning phase and did not compare the two videos. Further, participants' ratings of the quality of the explanation, comprehension, or perceived difficulty of the video did not differ significantly, and there was no significant difference in quality ratings in Experiment 1. Most importantly, it should be noted that learners' ratings of the quality of the videos was not associated with learning outcomes (retention: $r = .12$, $p = .132$, application: $r = .12$, $p = .152$).

General Discussion

Two experiments investigated the effects of an instructional video's setting on learning outcomes. The manipulation of setting was successful, as evidenced by the finding that learners rated the authentic setting to be more fitting for the learning materials than the neutral setting in both experiments. However, this did not reliably affect learning outcomes. Whereas Experiment 1 suggested that learning with a video that was shot in an authentic setting resulted in better retention than learning with an instructional video presenting the same contents in a neutral setting, this finding did not replicate in Experiment 2 with a larger sample size.[1] Further, across both experiments, there was no evidence for effects of setting on learners' capability to apply the learned contents nor on their assessments of the instructor's expertise. Hence, the current set of experiments does not provide reliable evidence in favor of the expertise hypothesis stating that an instructional video's setting has a positive effect on learning outcomes due to an increase of an instructor's purported expertise.

It is important to note that even though the authentic setting did not have a positive effect on learning, it also did not hamper learning; that is, the pattern of results contradicted the *distraction hypothesis* that would predict a negative effect of an authentic setting on learning outcomes and learners' self-reported distraction. Thus, although it does not seem to help, there also seems to be no harm in shooting instructional videos in a more detail-rich authentic context. Note though that the background did not include movement and that the authentic setting might have induced (more) distraction if that had been the case, because the onset of motion captures attention (Abrams & Christ, 2003). Thus, it is an open question for future research whether our findings generalize to videos shot in settings with more dynamic backgrounds.

Finally, there was no evidence for the *retrieval cue hypothesis* stating that the setting may constitute a contextual retrieval cue that facilitates performance on knowledge tests when the context from the learning phase is reinstated while testing; that is, there were no context effects in Experiment 2. An essential precondition for context effects to occur is that learners actually pay attention to context when encoding learning materials (Smith & Vela, 2001). Because learners rated the authentic setting to be more fitting for the presentation of learning materials than the neutral setting, they must have noticed the setting. However, there are potential boundary conditions (Smith & Vela, 2001) that may explain the lack of a context effect in Experiment 2.

First, context effects are smaller if learners engage in associative processing during the learning phase. In particular, it is assumed that this competes with the generation of unique associations between the context and the learning materials (Smith & Vela, 2001). Whereas learning materials that were used in previous studies demonstrating a context effect did not require associative processing, as these studies mainly used word lists, nonsense syllables, or faces (see Smith & Vela, 2001), the instructional videos that were used in the reported experiments did require associative processing. Consequently, learners may have built rich mental representations based on automatic associations between different elements of the learning materials (see Tibus, Heier, & Schwan, 2013) rather than relying on other contextual cues such as setting. If their mental representation was not enriched by the authentic context, it is not entirely surprising that learners did not benefit from reinstating it as a retrieval cue during the test phase.

A second potential explanation is that there was a short interval between learning and testing in Experiment 2 (i.e., an immediate posttest). Even though Smith and Vela (2001) still identified small context effects for short intervals between learning and testing (less than 5 min up to 1 day) in their meta-analysis, testing participants after longer intervals (1 day up to a week) resulted in medium-sized context effects. Hence, having a short interval between the learning and test phase in combination with learning materials that encourage spontaneous associations between differ-

[1] The ESCI macro (Cumming, 2012, ESCI software: www.thenewstatistics.com) was used to conduct a small-scale meta-analysis for the effect of setting during the learning phase on retention, including the two groups that were comparable across both experiments. This analysis revealed no significant overall effect, $d_{unbiased} = 0.30$, $p = 391$, 95% CI [−0.38, 0.97].

ent pieces of information may have reduced the probability of finding context effects.

A strength of this study is that, after finding a positive effect of an authentic setting on retention when testing two competing hypotheses in Experiment 1, a second experiment was conducted with a larger sample size to test whether the findings would replicate. Investigations of whether an initial exploratory finding would replicate are important because overinterpreting the reliability of effects observed in exploratory experiments may be considered a major factor contributing to the replication crisis in psychology (Open Science Collaboration, 2015). Direct replications allow for both the consolidation of exploratory findings as well as for the identification of false positives that may decrease the reliability and the reproducibility of the published research. Given the mixed findings across the two experiments reported in this article, the findings of Experiment 1 should not be seen as evidence in favor of shooting instructional videos in authentic settings.

However, even though the findings of Experiment 1 were not replicated, this set of experiments may inspire future research that identifies important boundary conditions for an effect of setting to occur (e.g., movement in the setting). Other directions for future research on potential effects of setting include a stronger integration of learning materials into the setting. Whereas we aimed at a clear separation between the fitting setting and the actual learning materials in the current experiments, it is an interesting question whether integrating learning materials into an authentic (and fitting) setting would affect learning. For example, with the materials used in these experiments, it would have been possible to integrate the plants that were used for demonstrations into the actual greenhouse setting.

Further, whereas the current experiments only compared an authentic (and fitting) setting to a neutral setting, future research could also address the effect of a setting that would not be a good fit for the learning materials (e.g., shooting the video about plants in a car service station). In particular, a mismatching setting may cause interfering thoughts and may thus distract learners' attention away from the actual learning materials. Moreover, rather than changing the on-screen setting, future experiments could vary the actual physical setting in which learners study instructional videos. Such research could yield valuable insights into the effects of following a lecture online (vs. in a lecture hall or on-the-job) and being tested online (vs. in the lecture hall or on-the-job).

In sum, the two experiments did not provide convincing evidence that shooting instructional videos in an authentic setting would improve learning outcomes compared to a neutral setting. Importantly, the (seductive) details in the authentic setting did not hamper learning either. Further research is necessary to identify potential boundary conditions for the effects of an instructional video's setting on learning outcomes. Such boundary conditions may, for example, include background movement in the setting.

References

Abrams, R. A., & Christ, S. E. (2003). Motion onset captures attention. *Psychological Science, 14,* 427–432. http://dx.doi.org/10.1111/1467-9280.01458

Adam, H., & Galinsky, A. D. (2012). Enclothed cognition. *Journal of Experimental Social Psychology, 48,* 918–925. http://dx.doi.org/10.1016/j.jesp.2012.02.008

Balch, W. R., Bowman, K., & Mohler, L. (1992). Music-dependent memory in immediate and delayed word recall. *Memory & Cognition, 20,* 21–28. http://dx.doi.org/10.3758/BF03208250

Beege, M., Schneider, S., Nebel, S., & Rey, G. D. (2017). Look into my eyes! Exploring the effect of addressing in educational videos. *Learning and Instruction, 49,* 113–120. http://dx.doi.org/10.1016/j.learninstruc.2017.01.004

Boekhout, P., van Gog, T., van de Wiel, M. W. J., Gerards-Last, D., & Geraets, J. (2010). Example-based learning: Effects of model expertise in relation to student expertise. *British Journal of Educational Psychology, 80,* 557–566. http://dx.doi.org/10.1348/000709910X497130

Boucheix, J.-M., Gauthier, P., Fontaine, J.-B., & Jaffeux, S. (2018). Mixed camera viewpoints improve learning medical hand procedure from video in nurse training? *Computers in Human Behavior, 89,* 418–429. http://dx.doi.org/10.1016/j.chb.2018.01.017

Cann, A., & Ross, D. A. (1989). Olfactory stimuli as context cues in human memory. *The American Journal of Psychology, 102,* 91–102. http://dx.doi.org/10.2307/1423118

Carney, R. N., & Levin, J. R. (2002). Pictorial illustrations still improve students' learning from text. *Educational Psychology Review, 14,* 5–26. http://dx.doi.org/10.1023/A:1013176309260

Cheng, J. T., Tracy, J. L., Foulsham, T., Kingstone, A., & Henrich, J. (2013). Two ways to the top: Evidence that dominance and prestige are distinct yet viable avenues to social rank and influence. *Journal of Personality and Social Psychology, 104,* 103–125. http://dx.doi.org/10.1037/a0030398

Choi, H. H., Van Merriënboer, J. J., & Paas, F. (2014). Effects of the physical environment on cognitive load and learning: Towards a new model of cognitive load. *Educational Psychology Review, 26,* 225–244. http://dx.doi.org/10.1007/s10648-014-9262-6

Cumming, G. (2012). *Understanding the new statistics: Effect sizes, confidence intervals, and meta-analysis*. New York, NY: Routledge.

de Koning, B., Hoogerheide, V., & Boucheix, J. M. (2018). Developments and trends in learning with instructional video. *Computers in Human Behavior, 89,* 395–398. http://dx.doi.org/10.1016/j.chb.2018.08.055

Doss, M. K., Picart, J. K., & Gallo, D. A. (2018). The dark side of context: Context reinstatement can distort memory. *Psychological Science, 29,* 914–925. http://dx.doi.org/10.1177/0956797617749534

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. http://dx.doi.org/10.3758/BRM.41.4.1149

Fiorella, L., Van Gog, T., Hoogerheide, V., & Mayer, R. E. (2017). It's all a matter of perspective: Viewing first-person video modeling examples promotes learning of an assembly task. *Journal of Educational Psychology, 109,* 653–665. http://dx.doi.org/10.1037/edu0000161

Garner, R., Brown, R., Sanders, S., & Menke, D. J. (1992). "Seductive details" and learning from text. In K. A. Renninger, S. Hidi, & A. Krapp (Eds.), *The role of interest in learning and development* (pp. 239–254). Hillsdale, NJ: Erlbaum.

Glick, S., Larsen, S., Johnson, C., & Branstiter, H. (2005). Evaluations of sexy women in low and high-status jobs. *Psychology of Women Quarterly, 29,* 389–395. http://dx.doi.org/10.1111/j.1471-6402.2005.00238.x

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

Harp, S. F., & Mayer, R. E. (1998). How seductive details do their damage: A theory of cognitive interest in science learning. *Journal of Educational Psychology, 90,* 414–434. http://dx.doi.org/10.1037/0022-0663.90.3.414

Hoogerheide, V., Loyens, S. M. M., Jadi, F., Vrins, A., & Van Gog, T. (2017). Testing the model-observer similarity hypothesis with text-based worked examples. *Educational Psychology, 37,* 112–127. http://dx.doi.org/10.1080/01443410.2015.1109609

Hoogerheide, V., Van Wermeskerken, M., Loyens, S. M. M., & Van Gog, T. (2016). Learning from video modeling examples: Content kept equal, adults are more effective models than peers. *Learning and Instruction, 44,* 22–30. http://dx.doi.org/10.1016/j.learninstruc.2016.02.004

Huff, M., Maurer, A. E., & Merkt, M. (2018). Producing gestures establishes a motor context for procedural learning tasks. *Learning and Instruction, 58,* 245–254. http://dx.doi.org/10.1016/j.learninstruc.2018.07.008

Jarvis, B. G. (2012). MediaLab (Version 2012) [Computer Software]. New York, NY: Empirisoft Corporation.

Kay, R. H. (2012). Exploring the use of video podcasts in education: A comprehensive review of the literature. *Computers in Human Behavior, 28,* 820–831. http://dx.doi.org/10.1016/j.chb.2012.01.011

Lachner, A., & Nückles, M. (2015). Bothered by abstractness or engaged by cohesion? Experts' explanations enhance novices' deep-learning. *Journal of Experimental Psychology: Applied, 21,* 101–115. http://dx.doi.org/10.1037/xap0000038

Lehman, S., Schraw, G., McCrudden, M. T., & Hartley, K. (2007). Processing and recall of seductive details in scientific text. *Contemporary Educational Psychology, 32,* 569–587. http://dx.doi.org/10.1016/j.cedpsych.2006.07.002

Lenzner, A., Schnotz, W., & Müller, A. (2013). The role of decorative pictures in learning. *Instructional Science, 41,* 811–831. http://dx.doi.org/10.1007/s11251-012-9256-z

Lüder, R. (2004). *Grundkurs Pflanzenbestimmung: Eine Praxisanleitung für Anfänger und Fortgeschrittene* [Basic course plant determination: A practical guide for beginners and advanced] (8th ed.). Wiebelsheim, Germany: Quelle & Meyer.

Mayer, R. (2014). *The Cambridge handbook of multimedia learning*. Cambridge, England: Cambridge University Press. http://dx.doi.org/10.1017/CBO9781139547369

Merkt, M., Ballmann, A., Felfeli, J., & Schwan, S. (2018). Pauses in educational videos: Testing the transience explanation against the structuring explanation. *Computers in Human Behavior, 89,* 399–410. http://dx.doi.org/10.1016/j.chb.2018.01.013

Merkt, M., Weigand, S., Heier, A., & Schwan, S. (2011). Learning with videos vs. learning with print: The role of interactive features. *Learning and Instruction, 21,* 687–704. http://dx.doi.org/10.1016/j.learninstruc.2011.03.004

Morris, C. D., Bransford, J. D., & Franks, J. J. (1977). Levels of processing versus transfer appropriate processing. *Journal of Verbal Learning & Verbal Behavior, 16,* 519–533. http://dx.doi.org/10.1016/S0022-5371(77)80016-9

Morris, T. L., Gorham, J., Cohen, S. H., & Huffman, D. (1996). Fashion in the classroom: Effects of attire on student perceptions of instructors in college classes. *Communication Education, 45,* 135–148. http://dx.doi.org/10.1080/03634529609379043

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

Open Science Collaboration. (2015). Estimating the reproducibility of psychological science. *Science, 349,* aac4716. http://dx.doi.org/10.1126/science.aac4716

Paas, F. (1992). Training strategies for attaining transfer of problem-solving skill in statistics: A cognitive-load approach. *Journal of Educational Psychology, 84,* 429–434. http://dx.doi.org/10.1037/0022-0663.84.4.429

Prenner, G., Bateman, R. M., & Rudall, P. J. (2010). Floral formulae updated for routine inclusion in formal taxonomic descriptions. *Taxon, 59,* 241–250. http://dx.doi.org/10.1002/tax.591022

Rand, G., & Wapner, S. (1967). Postural status as a factor in memory. *Journal of Verbal Learning & Verbal Behavior, 6,* 268–271. http://dx.doi.org/10.1016/S0022-5371(67)80107-5

Rey, G. D. (2012). How seductive are decorative elements in learning materials? *Journal of Educational Multimedia and Hypermedia, 21,* 257–283.

Rey, G. D. (2014). Seductive details and attention distraction—An eye tracker experiment. *Computers in Human Behavior, 32,* 133–144. http://dx.doi.org/10.1016/j.chb.2013.11.017

Ronse De Craene, L. P. (2010). *Floral diagrams: An aid to understanding flower morphology and evolution*. Cambridge, England: Cambridge University Press. http://dx.doi.org/10.1017/CBO9780511806711

Sanchez, C. A., & Wiley, J. (2006). An examination of the seductive details effect in terms of working memory capacity. *Memory & Cognition, 34,* 344–355. http://dx.doi.org/10.3758/BF03193412

Schwarz, N. (2007). Attitude construction: Evaluation in context. *Social Cognition, 25,* 638–656. http://dx.doi.org/10.1521/soco.2007.25.5.638

Serlio, S. (1991). Von den Szenen oder Schauplätzen. In K. Lazarowicz & C. Balme (Eds.), *Texte zur Theorie des Theaters* (pp. 409–416). Stuttgart, Germany: Reclam.

Smith, S. M., Handy, J. D., Angello, G., & Manzano, I. (2014). Effects of similarity on environmental context cueing. *Memory, 22,* 493–508. http://dx.doi.org/10.1080/09658211.2013.800553

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

Smith, S. M., & Sinha, A. K. (1987). *Effects of brief immersion in a flotation tank on memory and cognition* (Tech. Rep. No. CSCS-004). College Station, TX: Texas A&M University, Committee for the Study of Cognitive Science.

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

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

Tibus, M., Heier, A., & Schwan, S. (2013). Do films make you learn? Inference processes in expository film comprehension. *Journal of Educational Psychology, 105,* 329–340. http://dx.doi.org/10.1037/a0030818

Tulving, E., & Thomson, D. M. (1973). Encoding specificity and retrieval processes in episodic memory. *Psychological Review, 80,* 352–373. http://dx.doi.org/10.1037/h0020071

Received March 13, 2019
Revision received August 14, 2019
Accepted August 15, 2019 ■