ORIGINAL RESEARCH

Do prior knowledge, model-observer similarity and social comparison influence the effectiveness of eye movement modeling examples for supporting multimedia learning?

Marie-Christin Krebs[1,2] · Anne Schüler[1] · Katharina Scheiter[1,3]

Abstract

We investigated in an experiment with 180 university students the joint role of prior knowledge, alleged model competence, and social comparison orientation regarding the effectiveness of Eye Movement Modeling Examples (EMME) for supporting multimedia learning. EMME consisted of short videos with gaze replays of an instructed model demonstrating effective multimedia processing strategies. Participants were either instructed that the model in the EMME-videos was a successful learner (competent model) or another participant (peer model). Participants in a control condition received no EMME. Furthermore, we activated domain-relevant prior knowledge in half of the participants before watching the EMME. Against our expectations, we found no influence of either prior knowledge activation or model-observer similarity. As expected, our results indicated that EMME fostered multimedia learning. This was also supported by findings from small-scale meta-analyses that were conducted with the focus on the effect of EMME for multimedia learning and potential moderators of the effect. Moreover, results showed first evidence that social comparison orientation interacts with (alleged) model competence regarding the effectiveness of EMME. Further research is needed to follow up on the influence of individual factors as well as social cues on the effectiveness of EMME.

Keywords Multimedia learning · Eye movement modeling examples · Observational learning · Social comparison orientation · Prior knowledge activation

✉ Marie-Christin Krebs
    marie-christin.krebs@psychol.uni-giessen.de

    Anne Schüler
    a.schueler@iwm-tuebingen.de

    Katharina Scheiter
    k.scheiter@iwm-tuebingen.de

1   Leibniz-Institut für Wissensmedien, Schleichstrasse 6, 72076 Tübingen, Germany

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

3   University of Tübingen, Geschwister-Scholl-Platz, 72074 Tübingen, Germany

Introduction

Learning with illustrated text (i.e., multimedia) has been shown to be more effective than learning with text alone (Butcher, 2014). This 'multimedia effect' (Mayer, 2009) underlines the importance of illustrations for learning. However, learning with multimedia is often challenging for learners. According to the Cognitive Theory of Multimedia Learning (CTML; Mayer, 2009) effective multimedia learning requires learners to actively perform certain cognitive processing strategies. They first must select relevant information from both the text and the picture, and then organize the selected information in separate mental models of text and picture, and finally integrate the mental models derived from text and picture with each other. Empirically, picture processing as well as integration have been shown to be especially important for learning successfully from illustrated texts (e.g., Eitel et al., 2013; Hegarty & Just, 1993; Mason et al., 2017).

However, learners differ in their ability to apply the described-above processing strategies (e.g., Hannus & Hyönä, 1999; Hegarty & Just, 1993) with poor learners demonstrating less effective processing strategies. A possible explanation why learners fail in applying the necessary cognitive strategies during the learning process is that they lack the ability to regulate their learning concerning their choice of effective cognitive strategies (Scheiter et al., 2017).

Eye Movement Modeling Examples (EMME; Jarodzka et al., 2013) consist of recordings of the eye movements of a model during performing a task. These eye movements are overlaid onto the to-be-processed task material, so learners can see where, for how long, and in what order the model looks at the task material. As an instructional tool, these gaze displays are shown to novices to help them perform the same or similar tasks. For example, in the context of multimedia learning EMME are used to illustrate effective processing strategies for illustrated texts (e.g., Krebs et al., 2019; Mason et al., 2016, 2017; Scheiter et al., 2017). In the present study, the EMME displayed a model's gaze replay of performing effective processing strategies (e.g., inspecting the picture before reading the text, looking at corresponding picture-elements during text-reading) that had been derived from the literature (cf., Krebs et al., 2019; Scheiter et al., 2017). The model's gaze fixation was depicted as a white spotlight on an otherwise grey shaded page of the learning material (see Fig. 2 for an example of the learning material with and without EMME-guidance). In the present study, EMME-guidance was presented on the first four pages of the later learning material.

Although EMME have been shown to be effective for multimedia learning (e.g., Krebs et al., 2019; Mason et al., 2015, 2016, 2017; Scheiter et al., 2017), there are also empirical indications that not all learners benefit equally from EMME. For example, previous research indicated that there are cognitive factors such as domain knowledge or reading comprehension skills that can influence the effectiveness of EMME (Scheiter et al., 2017; Mason et al., 2016). Moreover, in line with the assumption that EMME as a special case of video-based modeling create a social learning situation (e.g., Jarodzka et al., 2013; Krebs et al., 2019), there is first empirical evidence that also social cues (e.g., model characteristics) that are provided in the introductory text can influence the effectiveness of EMME (Krebs et al., 2019). More specifically, the effectiveness of EMME has been shown to be more pronounced for learners with lower prior knowledge when the model was introduced as another learner in the same experiment [i.e., high (alleged) model-observer similarity; Krebs et al., 2019]. This result was attributed to the fact that modeling examples as a social learning situation encourage learners to engage in a social comparison with the

model (Hoogerheide et al., 2018). However, as there are individual differences regarding a person's tendency to compare oneself to others (i.e., social comparison orientation; Gibbons & Buunk, 1999), this might result in some people tending to engage more in social comparisons, which could influence the effect of (alleged) model-observer similarity on modeling examples.

Against this background, the aim of the study was to extend the research conducted by Krebs et al., (2019) by further investigating the relative influence of cognitive factors and social cues on the effectiveness of EMME. Additionally, the study aimed at investigating the role of social comparison orientation on the influence of social cues.

Using EMME for learning with illustrated texts

The effectiveness of EMME as instructional tool has been investigated in different areas and for different tasks (see Krebs et al., 2019, for an overview). Regarding learning with illustrated texts, EMME have been shown to be effective (Krebs et al., 2019; Mason et al., 2016, 2017; Scheiter et al., 2017). However, there are first empirical indications that there are factors which influence the effectiveness of EMME (e.g., Krebs et al., 2019; Mason et al., 2016; Scheiter et al., 2017). According to Krebs et al. (2019), there are different theoretical perspectives on why EMME are effective for multimedia learning. Importantly, these perspectives are not mutually exclusive, but are likely to complement each other. Based on these theoretical perspectives, potential factors can be derived that might influence the effectiveness of EMME (Krebs et al., 2019).

From an attentional perspective, EMME are effective because they highlight relevant information, which helps learners to select this information more quickly. Highlighting relevant information to guide learners' visual attention to relevant information is usually referred to as signaling principle, cuing principle, or attention-guiding principle (Richter et al., 2016; Schneider et al., 2018; Van Gog, 2014). Gaze cues can be considered a special form of attention-guiding cue (Gallagher-Mitchell et al., 2018; Litchfield et al., 2010). Thus, it can be assumed that gaze cues follow the same mechanisms. Moreover, it is presumed that synchronization processes are triggered by observing the model's eye movements (Jarodzka et al., 2013; Van Marlen et al., 2016), which establishes a form of joint attention and allows learners to infer which information is task relevant (e.g., Chisari et al., 2020; Van Gog et al., 2009). Accordingly, EMME are supposed to support learners in selecting relevant information more quickly. In line with this assumption, Chisari et al. (2020), for example, recently demonstrated that learners with EMME-guidance looked faster at referenced information, which was associated with higher learning outcomes.

From a (meta-) cognitive perspective, EMME are assumed to support learners in acquiring new cognitive processing strategies (e.g., Mason et al., 2015, 2016) and/or in better regulating already existing ones (e.g., Mason et al., 2017; Scheiter et al., 2017). This is based on the assumption that there is a close link between people's fixation location and their cognitive processes (eye-mind assumption; Just & Carpenter, 1980). In line with this assumption, previous research found empirical evidence that not only cognitive processes direct attentional processes, but also vice versa. For example, Grant and Spivey (2003) demonstrated that directing learners' attention to crucial areas improved reasoning in a diagram-based problem-solving task. Based on their results, Grant and Spivey (2003) concluded: "Although it may often seem that attention and eye movements are the result of cognitive processing, it may be that sometimes cognitive processing is the result of attention and eye movements" (p. 466). Accordingly, observing the dynamic eye movements of

the model might not only trigger perceptual imitation processes, but also (similar) cognitive processes. Against this background, it is assumed that observing the model progressively applying effective multimedia processing strategies to the learning material supports learners in generating or activating a cognitive representation (cf. cognitive schema; Sweller et al., 2011) of the use of these strategies (i.e., multimedia processing schema). More specifically, by observing the EMME-videos, learners perceive a vivid example on how they should apply these strategies themselves. Consequently, EMME can be regarded as a particular type of worked example (i.e., video modeling example; Jarodzka et al., 2013). Because learning from examples has proven to be a very effective way to acquire cognitive skills (Van Gog & Rummel, 2010), the effectiveness of EMME might also be attributed to their role as instructional tool.

Another explanation for the effectiveness of EMME goes beyond a mere perceptual or cognitive explanation. In contrast to other forms of attention-guiding methods, learners' attention is guided by showing the actual eye movements of a human model rather than an abstract cue. Using human eye movements to perceptually guide learners' attention is based on findings that individuals develop sensitivity to the communicative function of other peoples' gaze from an early age and use it as an important source of information (e.g., Frischen et al., 2007; Gallagher-Mitchell et al., 2018; Symons, et al., 2004). Hence, gaze cues might include an additional social dimension that is missing in other forms of attention-guiding cues.

However, the question arises of whether learners are able to interpret the pattern of the rather abstract gaze cue as human gaze behavior and therefore are able to perceive EMME as a social learning situation. Various findings from previous research speak in favor of this assumption. For instance, findings from Van Wermeskerken et al. (2018) indicated that people are capable of interpreting abstract cues as human eye movements as well as of making sense of dynamic eye movement patterns of other persons that are visualized by an abstract stimulus (i.e., moving circle). Furthermore, there is empirical evidence that ascribing anthropomorphic qualities to abstract stimuli (e.g., moving dots; Stanley et al., 2007), and agency-instructions (i.e., human agent vs. artificial agent; Stanley et al., 2010) can influence people's perception and processing of a stimulus. Consistent with these findings, Gobel et al. (2018) demonstrated that the belief in the social origin (i.e., human gaze) of a non-social cue (red dot) influenced participants' performance in a joined spatial cueing task: Participants who were informed that the red dot had a human origin (i.e., that the red dot represented where their partner had previously looked at on the screen), showed a more efficient visual search behavior than participants who were informed that the screen position of the red dot was generated randomly. Against this background, it can be argued that when learners are informed in the introductory text that the abstract gaze cues used in the EMME (e.g., moving dot; circle; spotlight) represent human gaze behavior, they are able to perceive even abstract cues as human eye movements and interpret them accordingly. Hence, in the case that learners perceive the rather abstract gaze cue as a social cue (i.e., human gaze cue), it is assumed that a social context (e.g., a social learning situation) emerges and thus also social factors come into play.

From a social-cognitive learning perspective, EMME can be regarded as a special case of video-based modeling (Jarodzka et al., 2013; Krebs et al., 2019). It is assumed that EMME induce a social learning situation in which a model conveys knowledge to a (novice) learner by externalizing his/her cognitive processing strategies via his/her gaze behavior. Moreover, the use of human eye movements as attention-guiding cue and the provision of social cues for learners in the introductory text are assumed to further increase the salience of a social situation, which in turn, fosters learning (social agency theory; Mayer

et al., 2003; Moreno et al., 2001). According to social agency theory (Mayer et al., 2003; Moreno et al., 2001), social cues prime learners to perceive a learning situation as a specific form of social interaction, thereby leading them to process the learning material more deeply and engage in more sense making processes which in turn foster learning outcome. Accordingly, there is empirical evidence that social cues such as onscreen agents displaying humanlike gesturing, movement, voice, eye contact, and facial expressions foster learning (e.g., Mayer, 2014; Moreno et al., 2001; Töpper et al., 2014; Wang et al., 2018). Against this background, it is assumed that the establishment of the EMME cue as 'human eye movements' potentially creates a social learning situation and thus may lead to deeper learning.

Previous research assumed that there are factors that can affect observational learning (Van Gog & Rummel, 2010). Although there are also mixed results, there is empirical evidence that for example learners' cognitive abilities such as prior (domain) knowledge, the perceived competence of a model and the similarity between model and observer (LeBel et al., 2018; Schunk, 1987) might affect learning from modeling examples. Model characteristics (e.g., age, competence/expertise, status) are assumed to be important for observational learning (Bandura, 1986), as, for example, (alleged) similarity offers an important source of information for gauging one's self-efficacy regarding a task (Schunk, 1987). Accordingly, learners in observational learning situations are considered to be more motivated to complete a task if they observe that similar others are successful in the same task because they assume that they can be equally successful (Schunk, 1987). However, if models are perceived to be too dissimilar, learners might have difficulties to identify themselves with the model and assume they cannot achieve the demonstrated behavior on their own (Schunk, 1987; Schunk & Hanson, 1985). Although according to social cognitive learning theory (Bandura, 1986) learners tend to pay more attention to effective models that have achieved good results with their behavior (Bandura, 1986; Schunk, 1987), it is therefore important that the perceived gap between the learner's competence and the model's competence is not too wide. Consequently, domain experts or models that are perceived by learners as too far out of their reach might be less effective. In contrast, models with only a relatively higher level of competence (e.g., peers who receive better marks in a test) might be more effective because learners perceive the model as more similar, which in turn fosters their self-efficacy beliefs and increases the likelihood that they will attempt to adopt the modeled behavior (Bandura, 1994; Schunk, 1987).

In line of reasoning that EMME can be regarded as a special form of (video) modeling examples (Jarodzka et al., 2013), it is possible that social cues such as (alleged) model-observer similarity also influence learning from EMME (Krebs et al., 2019). Recent research supports this assumption. Krebs et al. (2019) demonstrated that learners with less domain knowledge were able to benefit from EMME, but only as long as they perceived the model as a 'peer' learner. However, for learners with more domain knowledge there was no influence of (alleged) model-observer similarity. In this study, participants in the competent-model condition received the information that they would see the recorded eye movements of a learner who used successful strategies and therefore scored well on the knowledge test. In contrast, participants in the peer-model condition received the information that they would see the recorded eye movements of a learner who participated earlier in the experiment. Results revealed a significant interaction effect between domain knowledge and condition ($p = 0.039$, *adj. $R^2 = 0.10$*), indicating that only for learners with relatively low domain knowledge the description of the model mattered. This finding aligns with the assumption by Hoogerheide (2016) that model-observer similarity for learning from (video) modeling examples might be

especially important for learners with lower prior knowledge. Prior to the study by Krebs et al. (2019), results from Scheiter et al. (2017) suggested that only learners with higher levels of prior knowledge benefited from EMME. However, in contrast to Krebs et al. (2019), potential effects of (alleged) model-observer similarity were not investigated in this study.

Overall, research on the influence of model-observer similarity on observational learning led to mixed effects. There is empirical evidence that model-observer similarity (e.g., regarding gender or alleged competence) can be beneficial for academic or problem-solving tasks (e.g., Braaksma et al., 2002; Hartmann et al., 2020; Kim, 2007). In contrast, other studies did not find an effect of model-observer similarity (e.g., Groenendijk et al., 2013; Hoogerheide et al., 2016, 2017).

There are different explanations for the mixed findings regarding the effectiveness of model-observer similarity on observational learning. First, research on the effect of model-observer similarity on learning is usually performed with regard to different dimensions such as age, gender, or (alleged) competence. Hence, there might be differences in how (alleged) similarity regarding these different dimensions affect learning. Moreover, this also raises another problem, namely that it can be difficult to examine the different dimensions independently of each other. For instance, (alleged) similarity with regard to competence / expertise is usually difficult to investigate because other model characteristics such as age, gender or status are often linked to certain expertise expectations (e.g. task-appropriateness; Hoogerheide et al., 2018). Therefore, in addition to explicit information about the expertise of the model, other model characteristics might also influence participants' perception of the model's competence/expertise. For instance, Van Marlen et al. (2018) found contrary to the study by Krebs et al. (2019) that (alleged) expertise of the model had no significant influence on the learning outcome. In contrast to the study by Krebs et al. (2019), however, the model was not introduced as another learner, but as a teacher (domain expert teacher vs. other domain teacher). It is possible that the introduction of the model's status (i.e., being a teacher) already triggered subconsciously general expertise expectations, and thus participants experienced the gap in expertise as being too wide in both expertise conditions. The finding that in both expertise conditions the explanation that was given by the model was rated as high quality supports this assumption.

Against this background, when investigating (alleged) similarity with regard to competence/expertise, it is important that the participants do not receive any information about other model characteristics. A closer look at other studies in the EMME context show that participants often receive not only explicit information about the model's competence/expertise, but also implicit or explicit information about other model characteristics such as gender or status (e.g., Chisari et al., 2020; Jarodzka et al., 2013; Van Marlen et al., 2018). To investigate only the effect of (alleged) similarity in terms of competence/expertise, in our previous experiments as well as in the present one we provided no further information about other model characteristics in the introductory text or in the EMME-videos (cf. Krebs et al., 2019; Scheiter et al., 2017). Since the EMME-videos in the conditions with EMME-instruction were identical and without verbal explanations, it was assumed that participants in the present study could derive information about the model only from the description they received in the introductory text.

Another possible explanation for the mixed findings regarding the effectiveness of model-observer similarity on observational learning is that modeling encourages social comparison (Hoogerheide et al., 2018; Schunk & Hanson, 1985), and thus individual

differences such as learners' level of prior knowledge and / or their tendency to compare themselves to others might also play a role.

Influence of learner characteristics

Prior knowledge has been found to play an important role in learning from modeling examples (Bandura, 1986; Van Gog & Rummel, 2010). According to the social-cognitive learning theory, especially learners with higher cognitive skills and prior knowledge benefit from observational learning (Bandura, 1986). However, there are mixed findings with regard to the role of prior knowledge in learning from EMME. Scheiter et al. (2017) demonstrated in line with Bandura's assumption that only learners with relatively higher domain knowledge (i.e., compared to the sample's average domain knowledge) benefited from EMME instruction. Moreover, in the study by Krebs et al., (2019) learners with lower domain knowledge benefited from EMME, but only as long as (alleged) model-observer similarity was high. In contrast, there is research by Van Marlen et al. (2018) showing that a sample of low prior knowledge learners benefited from EMME (Experiment 2), whereas a sample of high prior knowledge learners did not benefit from any instructional support (Experiment 1). Furthermore, in a recent study, Chisari et al. (2020) observed no influence of prior knowledge (high vs. low) on the effectiveness of EMME.

It is possible that these contradictory findings are due to different definitions of the levels of prior knowledge. While Scheiter et al. (2017) and Krebs et al. (2019) referred to the level of domain prior knowledge in a relative sense (i.e., higher domain prior knowledge implied that it was higher compared to the mean level of prior knowledge in the respective study sample), Van Marlen et al. (2018) referred to the level of domain prior knowledge in a more absolute sense (i.e., higher domain prior knowledge implied that it was high with regard to the absolute test values). Furthermore, whereas in the experiments by Scheiter et al. (2017) and Krebs et al. (2019) there was still the possibility for performance improvement, this possibility did not exist in Experiment 1 of Van Marlen et al. (2018). In contrast, Chisari et al. (2020) manipulated high vs. low prior knowledge by varying whether participants had a pre-training before the learning phase (i.e., high prior knowledge) or not (i.e., low prior knowledge). In sum, based on the described empirical studies it is difficult to draw a conclusion about the appropriate level of prior knowledge to benefit from EMME instruction.

In accordance with the assumption that a higher level of prior knowledge might be beneficial to benefit from instructional support, there is empirical evidence that activating domain-relevant prior knowledge even of those learners with relatively little prior knowledge can foster learning performance (Wetzels et al., 2011). These effects are explained by assuming that learners are provided with a relevant context in which they can integrate new information (Wetzels et al., 2011). Accordingly, activating learners' domain-relevant prior knowledge might enable learners with less prior knowledge to benefit from EMME regardless of alleged model-observer similarity.

The importance of considering learners' level of prior knowledge in observational learning is also underlined by empirical findings that especially learners with less prior knowledge tend to engage more in social comparison (Buunk et al., 2003). Social comparison is assumed to be a fundamental process that occurs frequently and oftentimes automatically without the full awareness of the individual engaging in the social comparison (Buunk & Gibbons, 2007). According to Gibbons and Buunk (1999), individual

differences regarding a persons' tendency to compare oneself to others (i.e., social comparison) result in that some people tend to engage more in social comparisons and act more upon these comparisons than others (e.g., Corcoran et al., 2011). Based on findings that social comparisons can influence people's judgments, experiences, and behavior by prompting them to relate information, for instance, about other people's behavior, abilities, and success to themselves (Corcoran et al., 2011), it is possible that learners' social comparison orientation also influences learning from modeling examples and therefore the effectiveness of EMME as well.

The present study

The study aimed at investigating the role of model-observer similarity, prior knowledge, and social comparison orientation for the effectiveness of EMME. More specifically, we investigated whether differences in learners' social comparison orientation would explain differences in learners' ability to benefit from EMME. Moreover, we attempted to replicate the findings by Krebs et al. (2019) regarding the influence of (alleged) model-observer similarity, suggesting that especially learners with lower domain knowledge would benefit more from (alleged) model-similarity. In addition, we were interested whether we would be able to find a positive effect of EMME for learners with lower domain knowledge independently of model-observer similarity when activating domain-relevant prior knowledge before the learning phase.

We conducted an experiment with a 2×3 between-subjects design. The first factor 'domain-relevant prior knowledge' (PK) referred to the activation of learners' domain-relevant prior knowledge. Before the learning phase, half of the participants received a text containing domain-relevant information (i.e., information about cell biology) to activate their domain-relevant prior knowledge. The other half received a text containing domain-irrelevant information (i.e., information about the Mayan history). By reading the text and subsequently recalling as many facts as possible from the text, domain-relevant prior knowledge was assumed to be either activated (cell biology text) or not activated (Mayan history text).

The second factor 'modeling condition' was also experimentally varied between participants. Two-third of the participants received EMME directly before the learning phase. These EMME consisted of recorded eye movements of an instructed model demonstrating effective processing strategies on the first four pages of the to-be-processed learning material. Following the study of Krebs et al. (2019) the participants were given different information about the model before the EMME presentation (i.e., that the model was introduced as being either highly competent or a peer model) to manipulate the (alleged) model competence. Importantly, whereas the instructions differed between EMME conditions, the EMME were the same between conditions. The remaining participants did not receive EMME (control condition).

Hypotheses

Against the theoretical background and the questions derived from it, we formulated the hypotheses listed below (see Fig. 1). Please note that in order to make the following hypotheses easier to understand, we refer to domain knowledge as if the study design had
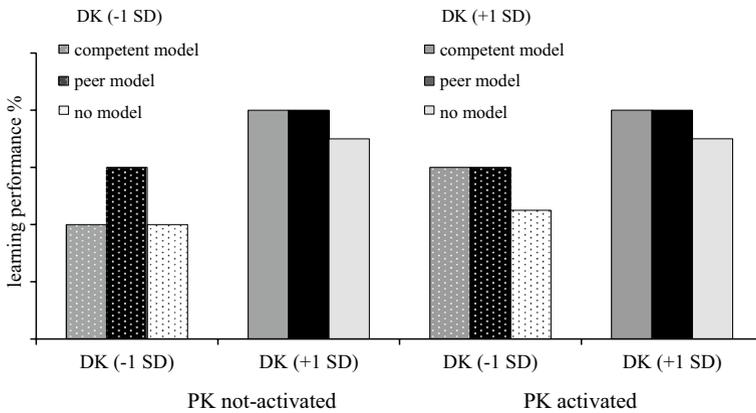
Fig. 1 Hypothesized relation between modeling condition, prior knowledge condition and domain knowledge. This figure illustrates the hypothesized relation between modeling condition (competent model vs. peer model vs. no model), prior knowledge condition (PK; with vs. without prior knowledge activation) and domain knowledge (DK; $+/-1$ SD of the continuous variable) against the backdrop of the model-observer similarity hypothesis

been based on discrete categories of low and high domain knowledge participants. However, in the analyses domain knowledge was included as a continuous predictor and the effect of domain knowledge was estimated by simple slope analyses at -1 SD and $+1$ SD of the variable (Aiken et al., 1991).

- Participants with more domain knowledge with EMME-support are expected to outperform participants without instructional support regardless of whether their prior knowledge had been previously activated (Hypothesis 1). Moreover, no significant difference in learning outcomes is expected between the two EMME conditions for learners with more domain knowledge (EMME effect irrespective of alleged model competence; Hypothesis 2).
- For participants with less domain knowledge and *without prior knowledge activation* we expected that model-observer similarity would influence the effectiveness of EMME. More specifically, in line with the model-observer similarity hypothesis and the findings by Krebs et al. (2019) we hypothesized that participants with less domain knowledge show higher learning performance when it is emphasized that the model participated in the same study as the learner (i.e., peer learner; high model-observer similarity) than when it is emphasized that the model is a "competent learner" (i.e., low model-observer similarity) or when they do not receive EMME-support (Hypothesis 3; see Fig. 1, left).
- For participants with less domain knowledge and *with prior knowledge activation*, however, we hypothesized based on positive effects of prior knowledge activation for learners with less domain knowledge (Wetzels et al., 2011) that they would benefit from EMME regardless of (alleged) model-observer similarity. Thus, we expected participants with less domain knowledge and with prior knowledge activation in both EMME conditions to outperform participants in the control condition (Hypothesis 4). Moreover, we did not expect a significant difference in learning outcomes between the two EMME conditions (Hypothesis 5) (see Fig. 1, right).

In addition to our hypotheses regarding the effect of prior knowledge activation and model observer-similarity for learners with more or less domain knowledge, we were interested in whether learner's tendency to compare themselves with others (i.e., social comparison orientation) would influence possible effects. As to our knowledge there are no previous studies investigating the construct in a similar context, we had no explicit hypotheses about the direction of possible effects. Therefore, we investigated the potential influence of social comparison orientation only exploratory.

Method

Participants and design

Initially, 180 students from a German university participated in the study. Six participants had to be excluded from data analyses due to technical and language problems. The remaining 174 students (122 female; $M = 22.48$ years, $SD = 3.62$) were enrolled in different university courses. The recommended sample size of 174 participants was determined by conducting a power analysis using G*Power 3.1 (Erdfelder et al., 2009) with the effect size *adjusted R^2* $= 0.10$ (derived from a prior study by Krebs et al., 2019), a power of 0.90, and alpha $= 0.05$. Due to the learning content (i.e., mitosis), students of biology, medicine, or related fields were excluded a priori from participating.

The design of the study was a 2×3 between-subjects design with domain-relevant prior knowledge (PK activated vs. PK not-activated) and modeling condition (competent vs. peer vs. control) as independent variables. Participants were randomly assigned to one of six conditions: PK not-activated/competent-model condition ($n = 30$), PK not-activated/peer-model condition ($n = 26$), PK not-activated/control condition ($n = 30$), PK activated/competent-model condition ($n = 27$), PK activated/peer-model condition ($n = 32$), PK activated/control condition ($n = 29$). Participation was voluntary and reimbursed with either 10 Euro or course credits. The reported study was approved by the local ethics board of the Leibniz-Institut für Wissensmedien (LEK 2015/019).

Materials

Activation of domain-relevant prior knowledge (PK)

To activate PK, half of the participants received domain-relevant information before the learning phase. The other half received domain-irrelevant information instead. Participants in the condition 'PK activated' received an explanatory text about cell biology with information regarding the history of its scientific discovery, description of cell processes and relevance for daily life. Importantly, the text did not contain the same information as conveyed during the learning phase (overall text length: 396 words). Based on the German curriculum for biology courses, which includes the learning topic 'cell division', we assumed that all participants had some prior knowledge that could be activated. Participants in the condition 'PK not-activated' received an explanatory text about the Mayan people with information about their history, culture and religion (overall text length: 400 words).

Learning material

Note, that since a similar method was used in Krebs et al. (2019), the material in the present study was the same as in the study by Scheiter et al. (2017) and Krebs et al. (2019). The material comprised an expository illustrated text that described relevant biological processes and principles on which mitosis is based. The learning content was distributed across 11 pages. The text was divided in semantically meaningful paragraphs with an overall length of 1113 words. Text lengths per page varied between 44 and 127 words. On each page, the text was displayed on the left, and the accompanying picture on the right-hand side. Text and picture were complementary with the text describing the processes during mitosis on a more abstract level and the pictures providing additional visual-spatial information about cell structures and processes during mitosis (see Fig. 2 for an example). To build a comprehensive mental model of the content, participants needed to process the text information as well as the picture information.

Based on previous research (Krebs et al., 2019; Scheiter et al., 2017), the learning material was divided into two parts. Part 1 of the learning material comprised four pages. In Part 1, the participants received basic information about important cell structures, chromosomes and the concept of DNA. Information from Part 1 was necessary to understand Part 2 of the learning material. Part 2 of the learning material comprised seven pages. These seven pages contained more detailed information on the basic concepts relevant to mitosis as well as on the phases of mitosis. First, participants were provided with a short overview about the whole process during mitosis. On the next page, the duplication of chromatin fibers and their development into chromosomes were described. Subsequently, the participants were provided with a description of the development of the mitotic spindle, the function of the equatorial plane, and the separation of sister chromatids. The final page referred to the segregation of daughter cells with genetically identical material. As in the previous study by Krebs et al. (2019), all learning outcome measures referred only to Part 2 of the learning material.

For participants in conditions without modeling, Part 1 was displayed without external visual guidance via EMME. For participants in conditions with modeling, on Part 1 of the learning material effective multimedia processing strategies were displayed via EMME.
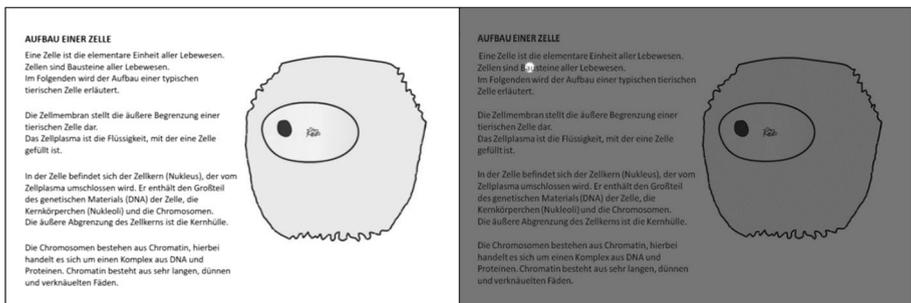

Fig. 2 Example of the learning material. This figure shows an example of the learning material (in German) explaining the terminology and the process in the text as well as showing spatial arrangements of the text-information in the picture. On the left, the first page of the learning material is displayed without EMME-guidance; on the right, the same page is displayed with EMME-guidance

The EMME that were displayed in the present study were identical to the EMME used in previous research (Krebs et al., 2019; Scheiter et al., 2017).

In order to generate the EMME, the model (a female student research assistant) was instructed on how to process the learning material and behave didactically by demonstrating each process as explicitly as possible (Krebs et al., 2019; Mason et al., 2016, 2017; Scheiter et al., 2017). First, the model read the title of the respective page and inspected the picture on the right-hand side of the page (construction of pictorial scaffold). This was followed by reading the text carefully and looking at corresponding elements of the picture (selection of relevant words and picture elements; text organization; picture organization). During reading the model switched between text and corresponding elements in the picture (integration). Before moving to the next page, the model took a final look at the picture (final picture inspection) and in some cases reread a text section (reaction to comprehension problems). Focus maps were used to visualize the model's eye movements on the learning material. A white spotlight on the otherwise grey shaded page represented a gaze fixation (see Fig. 2 for an example). The size of the white spotlight changed dynamically depending on the fixation duration (i.e., for longer fixations the white spotlight became larger). The average size of the spotlight corresponded approximately to one line in the text. The overall duration of the EMME videos was 382 s and varied between 52 and 115 s per page ($M = 95.5$ s). There was no verbal instruction accompanying the gaze displays.

We used a remote eye tracker (RED 250) from SensoMotoric Instruments (SMI; Teltow, Germany) with a sampling rate of 250 Hz and the iView X 2.2 and Experiment Center 3.3 software (SMI; Teltow, Germany) to record and prepare the eye movements of the model (cf. Scheiter et al., 2017).

Introductory text

In the introductory text, participants were informed that they would be able to learn at their own pace, but would not be able to go back to previous pages. Participants in the EMME-conditions were additionally instructed that they would see short videos with a learner's eye movements, which would be illustrated by a white spotlight on a grey shaded page. Moreover, they were informed that the size of the spotlight illustrates the fixation time (i.e., larger spots illustrating longer fixation times). Additionally, we manipulated (alleged) model- observer similarity by varying the instruction in the introductory text (cf. Krebs et al., 2019). Participants in the competent-model conditions were informed that they would see the recorded eye movements of a very successful learner who used learning strategies that are particularly effective for learning. To emphasize that the success of the model was no coincidence, we referred in the text not only to the success of the model at the end of the task, but also to the application of effective strategies. In contrast, participants in the peer-model conditions were informed that they would see the recorded eye movements of a learner who participated earlier in the experiment. Based on the theoretical assumption that people tend to project automatically their own experiences or attributes onto others—especially if they believe that they belong to their in-group (egocentrism; e.g., Krueger, 2003)—we emphasized that the other learner had previously participated in the same experiment. Since the participants were aware that they were participating in a learning experiment and since they did not receive any further information about the model, we expected that the participants would assume that the model would have similar performance prerequisites as they had themselves (See Fig. 1).

Measures

Participants' domain knowledge, social comparison orientation, and general self-efficacy expectations were assessed before the activation of learners' prior knowledge and the learning phase. After the learning phase, posttest performance was assessed as dependent variable. Furthermore, we examined participants' judgement of learning before as well as participants' interest and motivation after the posttest. The measures for participants' domain knowledge and the learning outcome measures were based on previous research (Krebs et al., 2019; Scheiter et al., 2017). All variables were recorded using the web-based survey software tool platform Qualtrics (http://www.qualtrics.com). Since Cronbach's alpha is not designed to accommodate guessing behavior (Paek, 2015), we used McDonald's omega (ω) as an estimate of reliability for domain knowledge scores and learning outcome measures. For calculating McDonald's omega, we used R software (R Core Team, 2017; version 3.6.2) with the psych package for reliability analyses (Revelle, 2019).

Domain knowledge

Participants' domain knowledge was assessed using a test of participants' prerequisite knowledge on cell division and a test of general scientific literacy.

Participants' prerequisite knowledge was assessed with 15 multiple-choice items. The items included questions about cell elements, genetics and mitosis (e.g., 'What are microtubules?' answer: 'Components of the spindle fibres'). Each item had four response options with one correct Answer. For each correct answer participants received one point, resulting in a maximum total score for prerequisite knowledge of 15 points. McDonald's omega was with a value of $\omega = 0.53$ rather low.

Participants' scientific literacy was measured by 24 items from the Life Sciences Scale of the Basic Scientific Literacy Test (Laugksch & Spargo, 1996). For each item participants had to rate statements about scientific processes or interrelationships between scientific concepts as either 'correct', 'incorrect', or 'unknown' (e.g., 'The elements that form the molecules of living beings are continuously recycled.' answer: 'correct'). For each correct answer, participants received one point, resulting in a maximum total score for scientific literacy of 24 points. McDonald's omega $\omega = 0.70$.

A significant correlation of both measures with posttest performance (all $p_s < 0.05$) indicated that both measures captured knowledge components that were relevant to the learning domain. Both measures were also significantly correlated with each other ($r = 0.31$, $p < 0.01$). Based on that, we used a combined measure of both measures referred to as domain knowledge. The combined domain knowledge score was calculated as the sum of the z-standardized scientific literacy and prior knowledge scores (cf. Krebs et al., 2019; Scheiter et al., 2017). McDonald's omega of the combined measure was $\omega = 0.72$.

Social comparison orientation

Social comparison orientation was assessed by 11 items of the Iowa-Netherlands Comparison Orientation Scale (INCOM; Gibbons & Buunk, 1999). For each item participants had to state on a 5-point Likert scale ('I disagree strongly'—'I agree strongly') if they tend to compare themselves with others (e.g., 'If I want to find out how well I have done something, I compare what I have done with how others have done'). Participants' social

comparison orientation score was calculated by summing the responses to each question. A higher score on the INCOM scale indicates a higher tendency to social comparison behaviors. Cronbach's alpha was $\alpha = 0.80$.

Learning outcome measures

Posttest performance measure comprised participants' recall and comprehension performance. It was assessed by text- and picture-based multiple-choice items as well as text- and picture-based forced-choice verification items. The multiple-choice items had four alternatives and one correct answer (e.g., 'What is not true about mitosis?' answer: 'Cytokinesis is the division of the cell nucleus'). For the forced-choice verification items participants had to state if the presented statements or pictures were either true or false (e.g., 'The kinetochore check whether the chromosomes are pulled from both sides with the same force.' Answer: 'incorrect'). For each correct answer participants received one point. For incorrect answers participants received zero points. As outcome variables, the percentage of correct answers was calculated for both measures.

Recall performance was assessed with three multiple-choice items and 17 forced-choice verification items (McDonald's omega $\omega = 0.63$). Comprehension performance was assessed with nine multiple-choice items and four forced-choice verification items (McDonald's omega $\omega = 0.63$). There was no time limit for answering the posttest. As in previous EMME-research (Krebs et al., 2019; Scheiter et al., 2017), all posttest items referred only to Part 2 of the learning material for which no EMME had been displayed to avoid direct influence of the external guidance.

Further measures

The following measures were assessed for exploratory reasons but were not included in the analyses: participants' general self-efficacy expectations (only at the beginning of the experiment), judgement of learning (JOL), interest and overall motivation, and intrinsic motivation (only in the EMME-conditions).

We assessed participants' general self-efficacy expectations at the beginning of the experiment based on the General Self-Efficacy Short Scale (ASKU; Beierlein et al., 2012). Participants rated the three items of the ASKU (e.g., 'In difficult situations I can rely on my abilities') on a five-point scale (doesn't apply at all– applies completely). For the general self-efficacy expectation score an averaged sum of the ratings was used. Cronbach's alpha was $\alpha = 0.77$.

Participants' JOL was assessed directly after the learning phase. Participants rated on a zero to 100 scale how confident they were that they could answer the following questions about cell division based on the knowledge they gained in the learning phase (cf., Schleinschok et al., 2017). For the JOL score, participants were required to rate their confidence as a percentage, where a higher percentage score indicated a higher confidence in their success. For the accuracy of their judgment, the difference between participants' overall test score and their JOL score was computed (cf., Schleinschok et al., 2017).

After the posttest, participants rated on a five-point scale (doesn't apply at all—applies completely) their interest and motivation during the learning phase using 10 items based on items from the situational interest measure (e.g., Rotgans & Schmidt, 2011). The items comprised for example participants' state ('I had fun during the learning phase'), their

interest in the learning content ('I want to know more about the learning content'), their perception of the learning process ('I was focused during the learning phase'), and their effort ('I made an effort to understand the learning content'). For the score an averaged sum of the ratings was used. Cronbach's alpha for the interest and motivation scale was $\alpha = 0.88$.

In the EMME-conditions, we assessed also participants' motivation with regard to the EMME-videos after the posttest. For this, we adapted 10 items from the intrinsic motivation short scale (KIM; Wilde et al., 2009). Here, participants rated on a five-point scale (doesn't apply at all—applies completely) their experiences with the EMME-videos (e.g., if they had fun watching the EMME-videos, if watching the EMME-videos was pleasant for them). For the EMME motivation score an averaged sum of the ratings was used. Cronbach's alpha for this scale was $\alpha = 0.79$. Means and standard deviations for the additional measures are displayed in Table 1.

Procedure

Participants were tested in small groups of two to five participants. Each participant was seated individually and was randomly assigned to one of six conditions. Each of the participants received paper-based written information on the experimental procedures. After having signed a consent form, participants answered computer-based demographic questions, the test on prior domain knowledge, the scientific literacy test, and the questionnaire on social-comparison orientation. Afterwards the prior knowledge activation took place. Participants were asked to read the upcoming information carefully and were instructed as follows: 'Please memorize important information from the text in a way that you are able to provide this information to a good friend later'. They were also informed that after four min they would be automatically forwarded to the next page. In the following, participants in the conditions with PK activated received information about cell biology. Participants in the conditions with PK not-activated received information about the Mayan people. After four min, participants were automatically forwarded to the next page, where they were asked to recall the information by describing the newly learned content to a good friend in four min time.

Then, participants received the instruction for the learning phase. All participants were informed that during the upcoming learning phase they would be able to learn at their own pace without a time limit. Moreover, they were informed that they only would be able to proceed to the next page after the word 'next' appeared. Furthermore, they were instructed that they would not be able to go back to previous pages. Moreover, participants in the EMME conditions were informed that they would see short videos with a learner's eye movements. Thereby, the eyes of the learner would be illustrated by a white spotlight on a grey shaded page with the size of the spotlight illustrating the learner's fixation time (i.e., larger spots illustrating longer fixation times).

Additionally, participants in the 'competent-model' conditions were informed as follows: 'On the first four pages of the learning material you will see the recorded eye movements of a very successful learner. This learner used learning strategies that are particularly effective for learning. Accordingly, this learner performed very well in the subsequent knowledge test.' In contrast, participants in the 'peer-model' conditions were informed as follows: 'On the first four pages of the learning material you will see the recorded eye movements of a learner who participated earlier in the experiment'. Importantly, the EMME-videos were, however, identical in all EMME conditions. After the

Table 1 Means and standard deviations (in parentheses) of additional measures

Measure	EMME (Competent-model)		EMME (Peer-model)		No modeling (control)	
	PK not-activated	PK activated	PK not-activated	PK activated	PK not-activated	PK activated
	$n=30$	$n=27$	$n=26$	$n=32$	$n=30$	$n=29$
General self-efficacy expectations (mean)	3.84 (0.53)	3.83 (0.42)	3.78 (0.84)	3.88 (0.65)	3.72 (0.58)	3.74 (0.45)
Judgement of learning (in percent)	60.40% (23.87)	59.41% (21.34)	59.35% (23.72)	58.66% (22.67)	49.34% (24.67)	46.33 (18.17)
Judgement of learning (accuracy)	−3.84 (20.35)	−5.80 (16.99)	−6.85 (18.80)	−4.13 (18.81)	−9.59 (25.07)	−9.22 (13.19)
Interest and motivation (mean)	3.27 (0.80)	3.38 (0.70)	3.46 (0.73)	3.40 (0.75)	3.25 (0.80)	3.97 (0.72)
Motivation with regard to instructional support tool (mean)	2.68 (0.62)	2.80 (0.76)	2.71 (0.74)	2.78 (0.68)	–	–

instruction, participants in the EMME conditions watched the EMME-videos on the first four pages (i.e., Part 1) of the learning material. Thereafter, they were shown the first four pages of the learning material again without EMME to give them the opportunity to study the material again at their own pace. For participants in the conditions without EMME, the first four pages (Part 1) of the learning phase were displayed directly without external guidance.

After Part 1 of the learning material, participants entered Part 2 of the learning material, which was identical for all conditions. To ensure a minimum learning time each page was displayed for 50 s before participants could decide on their own to continue by pressing the space bar. After the learning phase, participants' posttest performance was assessed. Finally, participants were debriefed and paid. Each session lasted about 60 min.

Data analyses

To investigate the effects of PK activation and alleged model competence on the effectiveness of EMME and the influence of learner domain knowledge and social comparison orientation results were analyzed in different steps.

First, we analyzed domain knowledge and social comparison orientation to determine whether the experimental conditions differed regarding these characteristics.

Subsequently, to test Hypotheses 1–5, we used two three-way interaction models (Hayes, 2018, PROCESS Model 3). In these models, modeling condition (competent model vs. peer model vs. no model) was entered as multicategorical focal predictor, PK activation as categorical moderator (activated vs. not-activated), the z-standardized domain knowledge as continuous moderator, and the respective learning outcome score (recall performance; comprehension performance) as dependent variable.

To exploratory investigate a potential influence of social comparison orientation, we conducted two multiple regression analyses with recall performance and comprehension performance as dependent variables, respectively. Modeling condition, PK activation, the z-standardized domain knowledge score and the z-standardized social comparison orientation score as well as the interaction terms for the respective two-way interactions and three-way interactions were entered as independent variables into the analyses. To keep these analyses as parsimonious as possible, interactions comprising both z-standardized continuous variables were excluded from the regression model, since they are difficult to interpret and were not backed up by hypotheses. To follow up on significant interactions, simple slopes were conducted to estimate the size of the effect of the modeling condition at different levels of participants' domain knowledge and/or social comparison orientation (Aiken et al., 1991). The effect was estimated at—1 SD for lower values, and at + 1 SD for higher values relative to the mean of the respective continuous variable. For all statistical analyses the α level was set to α = 0.05. *Adjusted R^2* values were reported as measure of effect size.

Results

To determine whether there were a priori differences between experimental conditions regarding participants' domain knowledge, social comparison orientation, general self-efficacy expectations, interest, motivation or learning time, we conducted regression analyses

with conditions as predictors. There were no significant differences among conditions with all $ps > 0.05$. Further results for interest, motivation or learning time are not included in the results section. Table 2 shows the correlations among learning outcome measures (i.e., recall performance and comprehension performance) and potential moderator (i.e., domain knowledge and social comparison orientation) as well as means and standard deviations at baseline.

Recall performance

Results for the three-way interaction model for recall performance as dependent variable indicated against our hypotheses no significant three-way interaction between modeling condition, PK activation and domain knowledge score, $F < 1$. Further, also in contrast to our hypotheses, neither PK activation nor domain knowledge interacted with the modeling condition. Thus, Hypotheses 1–5 were not supported by the results of the three-way interaction model.

For exploratory purposes, we conducted in a next step a multiple regression analysis with modeling condition, PK activation, domain knowledge score and social comparison orientation score as predictors. Results showed a significant effect of domain knowledge, $F (1, 156) = 8.12$, $p = 0.005$, *adj.* $R^2 = 0.08$, indicating that participants with relatively high domain knowledge ($M = 71.14\%$, $SE = 1.45$) showed a better recall performance than participants with relatively low domain knowledge ($M = 65.18\%$, $SE = 1.46$). Furthermore, there was a significant effect of modeling condition, $F (2, 156) = 4.54$, $p = 0.012$, *adj.* $R^2 = 0.08$. Pairwise comparisons revealed that participants in the competent-model condition, $b = 6.95$, $SE = 2.54$, $\beta = 0.242$, $p = 0.007$, 95% CI [1.94, 11.95], as well as in the peer-model condition, $b = 6.09$, $SE = 2.48$, $\beta = 0.213$, $p = 0.015$, 95% CI [1.19, 10.98] performed better than participants without EMME. There was, however, no significant difference between both EMME conditions, $b = -0.86$, $SE = 2.51$, $\beta = -0.030$, $p = 0.732$, 95% CI [-5.82, 4.10] (see Table 3). Results showed no effect of prior knowledge activation, $F < 1$, or social comparison orientation, $F < 1$. None of the two-way interactions were

Table 2 Correlations among learning outcome measures and potential moderators as well as means and standard deviations at baseline

Measure	1[a]	2[a]	3[b]	4[b]
1. Recall[a]	–			
2. Comprehension[a]	0.382^{***}	–		
3. Domain knowledge[b]	0.260^{**}	0.255^{**}	–	
4. Social comparison orientation[b]	-0.029	-0.008	0.069	–
M	68.19	52.56	0.00	0.00
SD	13.48	15.54	1.00	1.00

$N = 174$. Since domain knowledge and social comparison orientation were z-standardized, the mean value was 0 and the standard deviation 1 for both measures

$*p < 0.05$

$**p < 0.01$

$***p < 0.001$

[a]Percentage correct

[b]z-standardized

Table 3 Means and standard errors for the learning performance measures as a function of modeling condition

Measure	Competent model (n = 57)		Peer model (n = 58)		Control (n = 59)	
	M	SE	M	SE	M	SE
Recall (in %)	70.76	1.82	69.90	1.73	63.82	1.77
Comprehension (in %)	54.82	2.05	55.23	1.96	48.59	2.00

significant (all $Fs < 1$, except for the interaction between modeling condition and social comparison orientation, $F (2, 156) = 1.52$, $p = 0.221$, *adj.* $R^2 = 0.08$). Furthermore, none of the three-way interactions were significant (all $Fs < 1$).

Overall, also the results of the multiple regression analysis for recall performance did not support our hypotheses. Moreover, results did not indicate a significant influence of social comparison orientation. However, one could argue that the finding that participants in the EMME conditions outperformed participants without EMME with no significant influence of PK activation, domain knowledge or model description suggests that all learners—including learners with higher prior knowledge—benefitted from EMME. This would partially support Hypothesis 1. Moreover, as there was no difference between the two EMME conditions, but only an overall effect of modeling, the results are also partially in line with Hypothesis 2 (EMME effect irrespective of alleged model competence).

Comprehension performance

Against our hypothesis, results for the three-way interaction model for comprehension performance as dependent variable indicated no significant three-way interaction between modeling condition, PK activation and domain knowledge score, $F (2, 162) = 2.50$, $p = 0.085$ (See Fig. 3). Again, neither PK activation nor domain knowledge interacted with
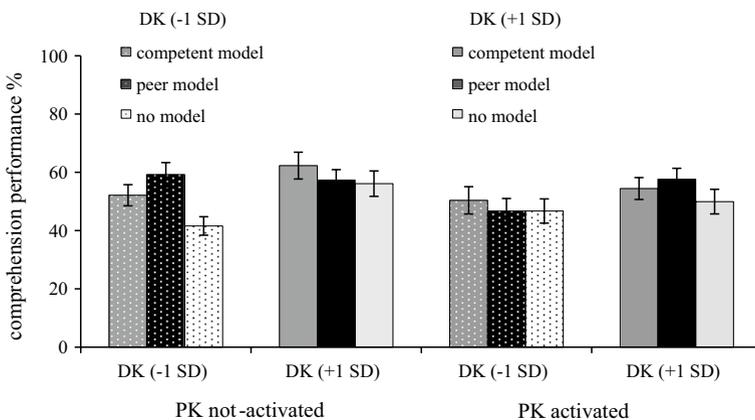

Fig. 3 Comprehension performance as a function of modeling condition, prior knowledge condition and domain knowledge. This figure illustrates comprehension performance in percent as a function of modeling condition (competent model vs. peer model vs. no model), prior knowledge condition (PK; with vs. without prior knowledge activation) and domain knowledge (DK; +/− 1 SD of the continuous variable)

the modeling condition. Thus, Hypotheses 1–5 were not supported by the results of the three-way interaction model.

For exploratory purposes, we conducted in a next step a multiple regression analysis with modeling condition, PK activation, domain knowledge score and social comparison orientation score as predictors. Results revealed a significant effect of domain knowledge, F (1, 156) = 8.38, p = 0.004, $adj.$ R^2 = 0.11, indicating that participants with relatively high domain knowledge (M = 56.30%, SE = 1.66) performed significantly better than participants with relatively low domain knowledge (M = 49.46%, SE = 1.65). There was no effect of PK activation condition, F (1, 156) = 2.71, p = 0.102, $adj.$ R^2 = 0.11, or social comparison orientation, F < 1. Furthermore, results showed a significant main effect of modeling condition, F (2, 156) = 3.46, p = 0.034, $adj.$ R^2 = 0.11. Pairwise comparisons revealed that participants in the competent-model condition (M = 54.82%, SE = 2.05), b = 6.22, SE = 2.86, β = 0.188, p = 0.031, 95% CI [0.56, 11.88], as well as in the peer-model condition (M = 55.23, SE = 1.96), b = 6.64, SE = 2.80, β = 0.202, p = 0.019, 95% CI [1.11, 12.17], performed better than participants without EMME (M = 48.59, SE = 2.00). There was no significant difference between the two EMME conditions, b = 0.28, SE = 2.88, β = 0.009, p = 0.923, 95% CI [-5.42, 5.98].

In addition, results indicated also a significant interaction between modeling condition and social comparison orientation, F (2, 156) = 3.58, p = 0.030, $adj.$ R^2 = 0.11. To investigate the effect of modeling condition at different levels of social comparison orientation, simple slope analyses were conducted. The omnibus effect of modeling condition for participants with relatively high social comparison orientation failed to be statistically significant, F (2, 156) = 3.03, p = 0.051, $adj.$ R^2 = 0.11. There was, however, a significant omnibus effect of modeling condition for participants with relatively low social comparison orientation, F (2, 156) = 3.63, p = 0.029, $adj.$ R^2 = 0.11. Results indicated that participants in the peer-model condition performed significantly better than participants in the control condition, b = 9.94, SE = 3.84, β = 0.302, p = 0.011, but not significantly better than participants in the competent-model condition, b = 7.58, SE = 4.06, β = 0.230, p = 0.064. Participants in the competent-model condition did not perform better than participants in the control condition, b = 2.36, SE = 4.10, β = 0.071, p = 0.567. None of the other two-way interactions were significant (all Fs < 1, except for the interaction between PK activation condition and social comparison orientation, F (1, 156) = 1.51, p = 0.221, $adj.$ R^2 = 0.11). Moreover, there were no significant three-way interactions (all Fs < 1, except for the three-way interaction between modeling condition, PK activation condition and domain knowledge, F (2, 156) = 2.50, p = 0.085, $adj.$ R^2 = 0.11).

Overall, also the results of the multiple regression analysis for comprehension performance did not support our hypotheses. Again, however, it could be argued that the results partially support Hypotheses 1 and 2, as learners generally benefited from EMME regardless of PK activation, model description or domain knowledge. For social comparison orientation, results indicated surprisingly an influence of modeling condition on the effectiveness of EMME for learners with lower social comparison orientation.

Posthoc small-scale meta-analyses

Investigating the influence of prior knowledge and (alleged) model-observer-similarity on the effectiveness of EMME on multimedia learning resulted in mixed results across studies. To take possible explanations such as lack of power of single studies, the possible influence of situational factors, and differences in the characteristics of the respective studies

into account, we conducted two post-hoc small-scale meta-analyses across studies that met the following criteria: EMME as instructional support for multimedia learning, EMME without additional audio commentary, a condition without modeling as control group, and learning outcome as continuous dependent variable.

We decided to run separate analyses for learning outcomes that referred more to recalling information from the learning material (i.e., recall performance), and learning outcomes that referred more to comprehension and knowledge transfer (comprehension/transfer performance). Accordingly, 27 effect sizes from six experiments were included in the analyses for recall performance (Krebs et al., 2019, sub.; Mason et al., 2015, 2016, 2017; Scheiter et al., 2017). Furthermore, 15 effect sizes from five experiments were included in the analyses for comprehension/transfer performance (Krebs et al., 2019, sub.; Mason et al., 2015, 2016, 2017).

Including multiple comparisons between intervention groups and a control group within an experiment into one meta-analysis, however, results in a unit-of-analysis error, as the effect sizes are correlated (Harrer et al., 2019). To reduce the unit-of-analysis error, we used a procedure proposed by Harrer et al. (2019): we split the number of participants in groups without EMME intervention according to the intervention groups (i.e., groups with EMME instruction) before calculating the effect sizes. Moreover, we split also the number of participants in the group with EMME intervention if more than one effect size was calculated for the respective group.

To provide a better estimation of the true EMME effect and the possible influence of learner and/or model characteristics, we conducted small-scale meta-analyses for recall performance and comprehension/transfer performance, respectively. To test the robustness of findings across multiple experiments, we used random-effects models (cf. Eitel & Scheiter, 2015).

For calculating the basic meta-analyses, we used the analysis SPSS script provided by Field and Gillett (2010). For recall performance [random-effects; $k=27$, $d=0.23$, 95% CI (0.08, 0.37), $SE=0.07$, $z=3.12$, $p=0.002$] as well as comprehension/transfer performance [random-effects; $k=15$, $d=0.28$, 95% CI (0.15, 0.41), $SE=0.07$, $z=4.21$, $p<0.001$], the basic meta-analyses revealed a small but significant effect of EMME instruction on multimedia learning.

In a next step, we investigated whether domain knowledge had an influence on the size of intervention effect for recall performance and/or comprehension/transfer performance. For calculating potential effects of domain knowledge as multicategorical moderator, we ran two separate meta-regression analyses for recall performance and comprehension/transfer performance, respectively. The categories for the level of domain knowledge were allocated with regard to the level of domain/prior knowledge in the respective studies: lower level of domain knowledge (-1 SD relative to the mean in the respective study sample) vs. medium level of domain knowledge (mean value in the respective study sample) vs. higher level of domain knowledge ($+1$ SD relative to the mean in the respective study sample). Against our expectations, results for recall performance, $F(2, 24)=1.01$, $SE=0.44$, $p=0.379$, as well as for comprehension/transfer performance, $F(2, 12)=2.01$, $SE=0.28$, $p=0.176$, $adj.$ $R^2=0.13$, showed no significant influence of domain knowledge as potential moderator on the effect of EMME on multimedia learning.

To test the influence of (alleged) model-observer similarity as another potential moderator, we ran two moderator variable analyses for recall performance and comprehension/transfer performance with (alleged) model-observer similarity as categorical variable using the analysis SPSS script provided by Field and Gillett (2010). Results for the moderator variable analyses showed against our expectations no significant effects for recall

performance, $\chi^2 = 0.02$; $p = 0.884$, or for comprehension/transfer performance, $\chi^2 < 0.01$; $p = 0.992$ across the studies.

Discussion

The aim of the present study was to investigate the influence of cognitive and social factors and a possible interaction between them on the effectiveness of EMME. For one, we sought to replicate the findings from Krebs et al. (2019) regarding the influence of (alleged) model-observer similarity on the effectiveness of EMME indicating that only learners with lower domain knowledge benefit from (alleged) model-similarity. For another, it was of interest whether prior knowledge activation would enable learners with lower prior knowledge to benefit from EMME independently of (alleged) model-observer similarity. Furthermore, it was investigated whether differences in learners' social comparison orientation would explain differences in learners' ability to benefit from EMME.

The results of our study did not directly confirm our hypotheses as the proposed three-way interaction model between modeling condition, prior knowledge activation and domain knowledge was not significant. However, some of the results reflected the underlying assumptions of our hypotheses. For example, results indicated that for learners with higher domain knowledge neither model-observer similarity nor prior knowledge activation influence the beneficial effect of EMME presentation. Therefore, it could be argued that these results at least partially support Hypotheses 1 and 2. Moreover, with regard to model-observer similarity these findings are in line with previous research by Krebs et al., (2019) and further underlines the premise that social cues (i.e., information about the model's competence) are not decisive for learners with higher domain knowledge. The results indicating that learners with higher domain knowledge did not benefit from prior knowledge activation were also in line with our expectations. For learners with higher domain knowledge we assumed that they already had more elaborate cognitive representations of learning content into which they could more easily integrate (new) information.

For learners with lower domain knowledge, however, we hypothesized in line with Wetzels et al. (2011) that they would benefit from prior knowledge activation, because they would be provided with a relevant context, which would help them to integrate new information. Further, we hypothesized that this would enable learners with lower domain knowledge to benefit from EMME even if the (alleged) model-observer similarity is low, because providing them with a relevant context would reduce the cognitive demands regarding the learning content. In contrast to our expectations, results revealed no beneficial effect of prior knowledge activation for learners with lower domain knowledge, thereby confirming neither Hypotheses 4 nor 5.

Moreover, contrary to our expectations and previous findings by Krebs et al. (2019) the results of the present study indicated no influence of (alleged) model-observer similarity for learners with lower domain knowledge. That is, in the present study all learners (irrespective of their prior knowledge) performed better when learning with EMME than without, an effect that was independent of the way the model's competence level had been characterized (not confirming Hypothesis 3). Thus, we were not able to replicate the findings by Krebs et al. (2019) that learners with lower prior knowledge only benefit from EMME when the (alleged) model-observer similarity is high.

A possible explanation for the mixed findings relates to the fact that compared to other studies that investigated the influence of social factors such as model-observer similarity

(e.g. Hoogerheide et al., 2016, 2018), the social context in the present study was rather weak as participants received information about the model only in the introductory text. Moreover, the information participants received was only about the model's performance (competent-model condition) or about the fact that the model participated earlier in the study (peer-model condition). Therefore, it is possible that the manipulation did not work properly because it was too subtle. If participants in the present study processed the information in the introductory text too inattentively, it is possible that they missed the information about the model. This might be problematic because in order for (alleged) similarity to be perceived, learners have to engage in social comparison processes (i.e., compare themselves with the model). However, as the present study was not conducted as an eye tracking experiment, there are no processing data available that might support this assumption. To investigate the effect of (alleged) model-observer similarity further, it is therefore important to ensure that enough social information is available to sufficiently activate a social context without revealing too much information about other model characteristics (e.g. age; gender).

The exploratory analyses regarding social comparison suggest that perhaps some learners are more sensitive to the description of the model in the introductory text. Results of the exploratory analysis for comprehension performance indicated an interaction between alleged model competence and learners' social comparison orientation on learners' comprehension performance. For learners with higher social comparison orientation, results revealed no significant influence of (alleged) model competence. Although learners with a lower social comparison orientation are less willing to compare themselves with others from the outset and thus the alleged competence of the model should not actually have an influence on learners' comprehension performance, our results pointed to the opposite. That is, learners with lower social comparison orientation only benefitted from EMME for their comprehension performance when they received the instruction that the model was a 'peer learner'. It is possible that when presenting a 'competent' model, a potential comparison situation was more salient and this automatically triggered internal defense mechanisms for these learners in order to avoid a comparison with this model. Maybe the 'peer' model was perceived as less threatening so the learners were actually able to benefit from the model. This idea is based on the assumption that if dissimilar models are perceived as too advanced, learners might assume that they cannot achieve the demonstrated behavior on their own (Schunk & Hanson, 1985). This, in turn, might result in learners feeling threatened, as they worry about a negative outcome.

Another possible explanation relates to the findings from Michinov and Michinov (2001) in the field of social psychology. They found that the presumed relation between similarity and attraction to others (e.g., Festinger, 1954) holds only for individuals with lower social comparison orientation. They explained this finding by the fact that individuals with higher social comparison orientation are in general more uncertain (Buunk & Gibbons, 2007), and thus (alleged) similarity is less important as they are attracted equally to similar and dissimilar others. In contrast, individuals with lower social comparison orientation tend to be less uncertain, and thus should be mainly attracted by similar others. Hence, it is possible, that also in the present study (alleged) similarity played only a role for learners with lower social comparison orientation because they felt more attracted by an alleged similar other one and, in turn, benefitted more from an alleged peer-model. Yet, these are only speculations referring to results from exploratory analyses that require further investigation in future studies. For this purpose, future studies should include, for example, measures for assessing the level of perceived similarity, for assessing learners' self-efficacy, and explicit measures for attraction.

Taken together, the results of the present study are in line with previous research that EMME are a beneficial instructional tool for multimedia learning (e.g., Mason et al., 2016, 2017). Our results revealed that overall learners with instructional support by EMME outperformed learners without instructional support with regard to their recall as well as comprehension performance—regardless of learners' level of domain knowledge, prior knowledge activation, and model-observer similarity. Additionally, we found first indications that learners' social comparison orientation might play a role for the influence of model descriptions. Nevertheless, it is important to note that the results for a social comparative orientation at this point are only exploratory and require further research before conclusions can be drawn as to whether they should be considered for the effectiveness of EMME for multimedia learning. With regard to a potential influence of domain knowledge and / or model-observer similarity, in contrast to previous studies we have not found that only learners with more domain knowledge could benefit from EMME (Scheiter et al., 2017), nor have we found evidence of an influence of (alleged) model-observer similarity on the effectiveness of EMME (Krebs et al., 2019).

Accordingly, results from the present study—with exception of the exploratory findings for the interaction between learners' social comparison orientation and modeling condition—did not support the assumption that an explanation for the effectiveness of EMME might go beyond a mere perceptual or cognitive explanation. To further examine this question on the basis of current research, we conducted post-hoc small-scale meta-analyses across previous studies that investigated EMME in the context of multimedia learning. The advantage of meta-analyses is that also possible explanations for mixed effects such as the lack of significance of individual studies, the possible influence of situational factors and differences in the characteristics of the respective studies can be taken into account. Results from the small-scale meta-analyses indicated an overall positive effect of EMME for multimedia learning. However, results did not indicate an influence of domain knowledge and/ or model-observer similarity on the effect of EMME on multimedia learning across studies neither for comprehension /transfer performance, nor for recall performance. Hence, while individual studies have revealed an influence of these moderators, currently there is no converging evidence that these moderators have a reliable effect on learning with EMME. However, it has to be kept in mind that there are yet very few studies that have investigated these moderators.

Overall, both the results of the present study and the results of the small-scale meta-analyses lead to the conclusion that, even if there are circumstances in which the effectiveness of EMME is influenced by model and/or learner characteristics, the effects will be unreliable at best. Consequently, these results rather suggest that the social aspect of EMME might play only a minor role for the effectiveness of EMME for multimedia learning. Hence, the question of whether considering the social context of EMME can add to the explanation to the effectiveness of EMME for multimedia learning is still open.

Limitations and implications for future studies

Although previous research found evidence that learner characteristics such as prior knowledge (Scheiter et al., 2017) and social cues such as model-observer-similarity (Krebs et al., 2019) moderate the effectiveness of EMME for multimedia learning, we were not able to replicate these findings. Moreover, providing learners with a domain-relevant prior knowledge background was not helpful to support learners with lower prior knowledge to benefit from EMME. This finding indicated that prior knowledge activation may not be the means

of choice to promote the effectiveness of EMME—at least not the type of prior knowledge activation we used in the present study. It is important, however, that we were able to demonstrate that EMME are suitable for multimedia learning not only for younger children (e.g., Mason et al., 2016), but also for adult learners (university students)—regardless of their domain knowledge or (alleged) model-observer similarity.

Yet, there are certain limitations to the results of the present study. On the one hand, all participants completed a pretest at the beginning of the study. The pretest itself contained domain-relevant prior knowledge questions that already could have activated their prior knowledge to a certain degree. On the other hand, all participants were randomly assigned to the activation conditions without taking their measured domain knowledge into account. Therefore, it is a possibility that learners' domain knowledge collided with the intended prior knowledge activation (e.g., learners' with high domain knowledge in the low prior knowledge activation condition). Furthermore, it is also possible that our method was not appropriate to activate learners' already existing prior knowledge about mitosis. Further, it could be that especially learners with lower already existing prior knowledge were not able to benefit from the prior knowledge activation intervention. Moreover, contrary to previous research we did not assess learners' information processing in the present study. That deprived us of the opportunity to investigate whether learners' domain knowledge, social comparison orientation or the respective model descriptions influence learners' information processing, albeit not their learning performance.

Another limitation of the present study is the reliability for our outcome measures (recall performance: $\omega = 0.63$; comprehension performance: $\omega = 0.63$). Lower reliability indicates that the respective test might not yield stable and consistent results, which would threaten the validity of the test. Accordingly, for future studies the reliability of the outcome measures should be considered more carefully.

Taking these limitations into account, further research is needed to follow up on the question of the influence of individual differences such as domain knowledge and / or social comparison orientation, as well as on the influence of social cues such as model-observer similarity on the effectiveness of EMME.

Conclusions

In sum, our results support the assumption that using EMME can foster multimedia learning, thereby further confirming EMME's function as an effective instructional tool. Although our findings do not provide direct evidence for the influence of social cues such as alleged model-observer similarity they provide first indications that individual differences such as social comparison orientation can influence the effect of model instruction on the effectiveness of EMME. In addition, our results show that the role of learners' prior knowledge remains unclear and requires further research. Taken together, further research is needed to follow up on the influence of individual factors as well as social cues on the effectiveness of EMME as instructional tool for multimedia learning.

Declarations

Conflicts of interest None.

Ethical approval The reported research received full ethical approval from the local ethics board of the Leibniz-Institut für Wissensmedien (LEK 2015/019).

Consent to participate We received written consent from participants and used a standard consent form to inform participants about the study. Participants' data was analyzed anonymously.

References

References marked with an asterisk indicate studies included in the small-scale meta-analyses

Aiken, L. S., Wesst, S. G., & Reno, R. R. (1991). *Multiple regression: Testing and interpreting interactions* (1st ed.). SAGE.

Bandura, A. (1986). *Social foundations of thought and action: A social cognitive theory*. Prentice-Hall.

Bandura, A. (1994). Self-efficacy. In V. S. Ramachaudran (Ed.), *Encyclopedia of human behavior* (Vol. 4, pp. 71–81). Academic Press. (Reprinted in H. Friedman [Ed.], *Encyclopedia of mental health*. Academic Press, 1998).

Beierlein, C., Kovaleva, A., Kemper, C. J., & Rammstedt, B. (2012). *ASKU - Allgemeine Selbstwirksamkeit Kurzskala [Fragebogen]*. In Leibniz-Zentrum für Psychologische Information und Dokumentation (ZPID) (Hrsg.), Elektronisches Testarchiv (PSYNDEX Tests-Nr. 9006490). Trier: ZPID. https://doi.org/10.23668/psycharchives.418.

Braaksma, M. A. H., Rijlaarsdam, G., & Van den Bergh, H. (2002). Observational learning and the effects of model-observer similarity. *Journal of Educational Psychology, 94*, 405–415. https://doi.org/10.1037/0022-0663.94.2.405

Butcher, K. R. (2014). The multimedia principle. In R. E. Mayer (Ed.), *The Cambridge handbook of multimedia learning* (2nd ed., pp. 174–205). Cambridge University Press. https://doi.org/10.1017/CBO9781139547369.010

Buunk, A. P., & Gibbons, F. X. (2007). Social comparison: The end of a theory and the emergence of a field. *Organizational Behavior and Human Decision Processes, 102*, 3–21. https://doi.org/10.1016/j.obhdp.2006.09.007

Buunk, B. P., Zurriaga, R., Gonzalez-Roma, V., & Subirats, M. (2003). Engaging in upward and downward comparisons as a determinant of relative deprivation at work: A longitudinal study. *Journal of Vocational Behavior, 62*, 370–388. https://doi.org/10.1016/S0001-8791(02)00015-5

Chisari, L. B., Mockevičiūtė, A., Ruitenburg, S. K., Van Vemde, L., Kok, E. M., & Van Gog, T. (2020). Effects of prior knowledge and joint attention on learning from eye movement modelling examples. *Journal of Computer Assisted Learning*. https://doi.org/10.1111/jcal.12428

Corcoran, K., Crusius, J., & Mussweiler, T. (2011). Social comparison: Motives, standards, and mechanisms. In D. Chadee (Ed.), *Theories in social psychology* (pp. 119–139). Wiley-Blackwell.

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

Eitel, A., Scheiter, K., & Schüler, A. (2013). How inspecting a picture affects processing of text in multimedia learning. *Applied Cognitive Psychology, 27*, 451–461. https://doi.org/10.1002/acp.2922

Erdfelder, E., Faul, F., 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

Festinger, L. (1954). A theory of social comparison processes. *Human Relations, 7*, 117–140. https://doi.org/10.1177/001872675400700202

Field, A. P., & Gillett, R. (2010). How to do a meta-analysis. *British Journal of Mathematical and Statistical Psychology, 63*, 665–694.

Frischen, A., Bayliss, A. P., & Tipper, S. P. (2007). Gaze cueing of attention: Visual attention, social cognition, and individual differences. *Psychological Bulletin, 133*, 694–724. https://doi.org/10.1037/0033-2909.133.4.694

Gallagher-Mitchell, T., Simms, V., & Litchfield, D. (2018). Learning from where 'eye' remotely look or point: Impact on number line estimation error in adults. *Quarterly Journal of Experimental Psychology, 71*, 1526–1534. https://doi.org/10.1080/17470218.2017.1335335

Gibbons, F. X., & Buunk, B. P. (1999). Individual differences in social comparison: Development of a scale of social comparison orientation. *Journal of Personality and Social Psychology, 76*, 129–142. https://doi.org/10.1037/0022-3514.76.1.129

Gobel, M. S., Tufft, M. R. A., & Richardson, D. C. (2018). Social beliefs and visual attention: How the social relevance of a cue influences spatial orienting. *Cognitive Science*. https://doi.org/10.1111/cogs.12529

Grant, E. R., & Spivey, M. J. (2003). Eye movements and problem solving: Guiding attention guides thought. *Psychological Science, 14*, 462–466. https://doi.org/10.1111/1467-9280.02454

Groenendijk, T., Janssen, T., Rijlaarsdam, G., & Van den Bergh, H. (2013). The effect of observational learning on students' performance, processes, and motivation in two creative domains. *British Journal of Educational Psychology, 83*, 3–28. https://doi.org/10.1111/j.2044-8279.2011.02052.x

Hannus, M., & Hyönä, J. (1999). Utilization of illustrations during learning of science textbook passages among low- and high-ability children. *Contemporary Educational Psychology, 24*, 95–123. https://doi.org/10.1006/ceps.1998.0987

Harrer, M., Cuijpers, P., Furukawa, T. A., & Ebert, D. D. (2019). Doing meta-analysis in R: A hands-on guide. PROTECT Lab Erlangen. https://doi.org/10.5281/zenodo.2551803.

Hartmann, C., Van Gog, T., & Rummel, N. (2020). Do examples of failure effectively prepare students for learning from subsequent instruction? *Applied Cognitive Psychology, 34*, 879–889. https://doi.org/10.1002/acp.3651

Hayes, A. F. (2018). *Introduction to mediation, moderation, and conditional process analysis: A regression-based approach* (2nd ed.). Guilford Press.

Hegarty, M., & Just, M. A. (1993). Constructing mental models of machines from text and diagrams. *Journal of Memory and Language, 32*, 717–742. https://doi.org/10.1006/jmla.1993.1036

Hoogerheide, V. (2016, October 20). *Effects of observing and creating video modeling examples on cognitive and motivational aspects of learning*. Erasmus University Rotterdam. Retrieved August 31, 2018, from https://repub.eur.nl/pub/93239/Hoogerheide-proefschrift-DEF.PDF.

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. https://doi.org/10.1080/01443410.2015.1109609

Hoogerheide, V., Loyens, S. M., & Van Gog, T. (2016). Learning from video modeling examples: Does gender matter? *Instructional Science, 44*, 69–86. https://doi.org/10.1007/s11251-015-9360-y

Hoogerheide, V., Van Wermeskerken, M., Van Nassau, H., & Van Gog, T. (2018). Model-observer similarity and task-appropriateness in learning from video modeling examples: Do model and student gender affect test performance, self-efficacy, and perceived competence? *Computers in Human Behavior, 89*, 457–464. https://doi.org/10.1016/j.chb.2017.11.012

Jarodzka, H., Van Gog, T., Dorr, M., Scheiter, K., & Gerjets, P. (2013). Learning to see: Guiding students' attention via a model's eye movements fosters learning. *Learning and Instruction, 25*, 62–70. https://doi.org/10.1016/j.learninstruc.2012.11.004

Just, M. A., & Carpenter, P. A. (1980). A theory of reading: From eye fixations to comprehension. *Psychological Review, 87*, 329–354. https://doi.org/10.1037/0033-295X.87.4.329

Kim, Y. (2007). Desirable characteristics of learning companions. *International Journal of Artificial Intelligence in Education*, *17*, 371–388. Retrieved November 11, 2016, from http://digitalcommons.usu.edu/itls_facpub/62/.

*Krebs, M.-C., Schüler, A., & Scheiter, K. (2019). Just follow my eyes: The influence of model-observer similarity on Eye Movement Modeling Examples. *Learning and Instruction*, *61*, 126–137. https://doi.org/10.1016/j.learninstruc.2018.10.005.

Krueger, J. I. (2003). Return of the ego–Self-referent information as a filter for social prediction: Comment on Karniol. *Psychological Review, 110*, 585–590. https://doi.org/10.1037/0033-295X.110.3.585

Laugksch, R. C., & Spargo, P. E. (1996). Construction of a paper-and-pencil test of basic scientific literacy based on selected literacy goals recommended by the American association for the advancement of science. *Public Understanding of Science, 5*, 331–359. https://doi.org/10.1088/0963-6625/5/4/003

LeBel, M. E., Haverstock, J., Cristancho, S., van Eimeren, L., & Buckingham, G. (2018). Observational learning during simulation-based training in arthroscopy: Is it useful to novices? *Journal of Surgical Education, 75*, 222–230. https://doi.org/10.1016/j.jsurg.2017.06.005

Litchfield, D., Ball, L. J., Donovan, T., Manning, D., & Crawford, T. J. (2010). Viewing another person's eye movements improves identification of pulmonary nodules in chest x-ray inspection. *Journal of Experimental Psychology, 16*, 251–262. https://doi.org/10.1037/a0020082

*Mason, L., Pluchino, P., & Tornatora, M. C. (2015). Eye-movement modeling of integrative reading of an illustrated text: Effects on processing and learning. *Contemporary Educational Psychology, 41*, 172–187. https://doi.org/10.1016/j.cedpsych.2015.01.004

*Mason, L., Pluchino, P., & Tornatora, M. C. (2016). Using eye-tracking technology as an indirect instruction tool to improve text and picture processing and learning. *British Journal of Educational Technology, 47*, 1083–1095. https://doi.org/10.1111/bjet.12271

*Mason, L., Scheiter, K., & Tornatora, M. C. (2017). Using eye movements to model the sequence of text-picture processing for multimedia comprehension. *Journal of Computer Assisted Learning, 33*, 443–460. https://doi.org/10.1111/jcal.12191

Mayer, R. E. (2009). *Multimedia learning* (2nd ed.). Cambridge University Press.

Mayer, R. E. (2014). Principles based on social cues in multimedia learning: Personalization, voice, image, and embodiment principles. In R. E. Mayer (Ed.), *The Cambridge handbook of multimedia learning* (2nd ed., pp. 345–368). Cambridge University Press. https://doi.org/10.1017/CBO9781139547369.017

Mayer, R. E., Sobko, K., & Mautone, P. D. (2003). Social cues in multimedia learning: Role of speaker's voice. *Journal of Educational Psychology, 95*, 419–425. https://doi.org/10.1037/0022-0663.95.2.419

Michinov, E., & Michinov, N. (2001). The similarity hypothesis: A test of the moderating role of social comparison orientation. *European Journal of Social Psychology, 31*, 549–555. https://doi.org/10.1002/ejsp.78

Moreno, R., Mayer, R. E., Spires, H. A., & Lester, J. C. (2001). The case for social agency in computer-based teaching: Do students learn more deeply when they interact with animated pedagogical agents? *Cognition and Instruction, 19*, 177–213. https://doi.org/10.1207/S1532690XCI1902

Paek, I. (2015). An investigation of the impact of guessing on coefficient α and reliability. *Applied Psychological Measurement, 39*, 264–277. https://doi.org/10.1177/0146621614559516

Revelle, W. R. (2019). psych: Procedures for psychological, psychometric, and personality research. *R Package, 1*(8), 4.

Richter, J., Scheiter, K., & Eitel, A. (2016). Signaling text-picture relations in multimedia learning: A comprehensive meta-analysis. *Educational Research Review, 17*, 19–36. https://doi.org/10.1016/j.edurev.2015.12.003

Rotgans, J. I., & Schmidt, H. G. (2011). Situational interest and academic achievement in the active-learning classroom. *Learning and Instruction, 21*, 58–67. https://doi.org/10.1016/j.learninstruc.2009.11.001

*Scheiter, K., Schubert, C., & Schüler, A. (2017). Self-regulated learning from illustrated text: Eye movement modelling to support use and regulation of cognitive processes during learning from multimedia. *British Journal of Educational Psychology*, *88*, 80–94. https://doi.org/10.1111/bjep.12175

Schleinschok, K., Eitel, A., & Scheiter, K. (2017). Do drawing tasks improve monitoring and control during learning from text? *Learning and Instruction, 51*, 10–25. https://doi.org/10.1016/j.learninstruc.2017.02.002

Schneider, S., Beege, M., Nebel, S., & Rey, G. D. (2018). A meta-analysis of how signaling affects learning with media. *Educational Research Review, 23*, 1–24. https://doi.org/10.1016/j.edurev.2017.11.001

Schunk, D. H. (1987). Peer models and children's behavioral change. *Review of Educational Research, 57*, 149–174. https://doi.org/10.3102/00346543057002149

Schunk, D. H., & Hanson, A. R. (1985). Peer models: Influence on children's self-efficacy and achievement. *Journal of Educational Psychology, 77*, 313–322.

Stanley, J., Gowen, E., & Miall, R. C. (2007). Effects of agency on movement interference during observation of a moving dot stimulus. *Journal of Experimental Psychology, 33*, 915–926. https://doi.org/10.1037/0096-1523.33.4.915

Stanley, J., Gowen, E., & Miall, R. C. (2010). How instructions modify perception: An fMRI study investigating brain areas involved in attributing human agency. *NeuroImage, 52*, 389–400. https://doi.org/10.1016/j.neuroimage.2010.04.025

Sweller, J., Ayres, P., & Kalyuga, S. (2011). *Cognitive load theory*. Springer. https://doi.org/10.1007/978-1-4419-8126-4

Symons, L. A., Lee, K., Cedrone, C. C., & Nishimura, M. (2004). What are you looking at? Acuity for triadic eye gaze. *The Journal of General Psychology, 131*, 451–469.

R Core Team (2017). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/.

Töpper, J., Glaser, M., & Schwan, S. (2014). Extending social cue based principles of multimedia learning beyond their immediate effects. *Learning and Instruction, 29*, 10–20. https://doi.org/10.1016/j.learninstruc.2013.07.002

Van Gog, T. (2014). The signaling (or cueing) principle in multimedia learning. In R. E. Mayer (Ed.), *The Cambridge handbook of multimedia learning* (2nd ed., pp. 263–278). Cambridge University Press. https://doi.org/10.1017/CBO9781139547369.014

Van Gog, T., Jarodzka, H., Scheiter, K., Gerjets, P., & Paas, F. (2009). Attention guidance during example study via the model's eye movements. *Computers in Human Behavior, 25*, 785–791. https://doi.org/10.1016/j.chb.2009.02.007

Van Gog, T., & Rummel, N. (2010). Example-based learning: Integrating cognitive and social-cognitive research perspectives. *Educational Psychology Review, 22*, 155–174. https://doi.org/10.1007/s10648-010-9134-7

Van Marlen, T., Van Wermeskerken, M., Jarodzka, H., & Van Gog, T. (2016). Showing a model's eye movements in examples does not improve learning of problem-solving tasks. *Computers in Human Behavior, 65*, 448–459. https://doi.org/10.1016/j.chb.2016.08.041

Van Marlen, T., Van Wermeskerken, M., Jarodzka, H., & Van Gog, T. (2018). Effectiveness of eye movement modeling examples in problem solving: The role of verbal ambiguity and prior knowledge. *Learning and Instruction, 58*, 274–283. https://doi.org/10.1016/j.learninstruc.2018.07.005

Van Wermeskerken, M., Litchfield, D., & Van Gog, T. (2018). What am I looking at? Interpreting dynamic and static gaze displays. *Cognitive Science, 42*, 220–252. https://doi.org/10.1111/cogs.12484

Wang, F., Li, W., Mayer, R. E., & Liu, H. (2018). Animated pedagogical agents as aids in multimedia learning: Effects on eye-fixations during learning and learning outcomes. *Journal of Educational Psychology, 110*, 250–268. https://doi.org/10.1037/edu0000221

Wetzels, S. A. J., Kester, L., & Van Merriënboer, J. J. G. (2011). Adapting prior knowledge activation: Mobilisation, perspective taking, and learners' prior knowledge. *Computers in Human Behavior, 27*, 16–21. https://doi.org/10.1016/j.chb.2010.05.004

Wilde, M., Bätz, K., Kovaleva, A., & Urhahne, D. (2009). Testing a short scale of intrinsic motivation. *Zeitschrift Für Didaktik Der Naturwissenschaften, 15*, 31–45.