



# Effects of related decorative pictures on learning and metacognition

Demian Scherer<sup>1</sup> · Annika Verkühlen<sup>1</sup> · Stephan Dutke<sup>1</sup>

Received: 10 March 2022 / Accepted: 4 January 2023 / Published online: 31 January 2023  
© The Author(s) 2023

## Abstract

Research suggests that explanatory pictures support learning, whereas pictures that distract processing resources from the main ideas of a text may impair learning and are considered as seductive illustrations. However, non-explanatory pictures that are related to the text and that do not tempt readers to focus illustrations more than the text's main ideas might improve learning through spreading activation processes during learning. This effect might be reflected in memory measures or in metacognitive measures. Therefore, in two experiments, we tested related decorative pictures for positive effects on knowledge and transfer performance and metacognitive monitoring. Results indicate positive effects of related pictures on knowledge acquisition (Experiment 1) and metacognitive monitoring (Experiment 2). In neither experiment, related pictures enhanced transfer performance or interest (measured as triggered and maintained situational interest). This pattern of results can be explained by automatic mutual facilitation of related materials based on spreading activation processes.

**Keywords** Decorative pictures · Seductive details · Metacognition · Multimedia learning · Interest

## Introduction

The well-established integrated model of text and picture comprehension (Schnotz, 2022; Schnotz & Bannert, 2003) and the cognitive theory of multimedia learning (Mayer, 2022) suggest that explanatory illustrations help readers to understand expository text. Numerous empirical studies support these theories (e.g., Leopold et al., 2019; Schweppe et al., 2015, for reviews, see Butcher et al., 2014; Levie & Lentz, 1982; Mayer, 2022). There is also a tradition of research on seductive illustrations showing detrimental effects of irrelevant illustrations on knowledge acquisition and transfer performance (for meta-analyses, see Rey, 2012; Sundararajan & Adesope, 2020). This pattern of results suggests that explanatory pictures support, but non-explanatory pictures hinder learning. However, pictures can

---

✉ Demian Scherer  
demian.scherer@uni-muenster.de

<sup>1</sup> University of Münster, Münster, Germany

serve numerous other functions beyond explaining the text's content and distracting attention away from it (e.g., Schneider et al., 2016).

In this study, we focus on *related* decorative pictures. Decorative pictures make learning materials more aesthetic or appealing, however, they do not transmit relevant information about the learning content (Lenzner et al., 2013). Imagine, for example, a text describing how to sort laundry before using a washing machine and imagine this text is decorated with pictures showing an open wardrobe with clothes and an empty washing machine. These depictions are (a) representational as they depict environments, and objects familiar to the recipient, (b) decorative as they have the function to make the outer appearance of the text more appealing or attractive, (c) non-explanative as they do not help to understand how clothes should be sorted before washing—the topic described in the text, (d) related to the learning contents as the pictures, for example in this case, refer to an environment in which clothes are often located. In the following, we refer to pictures with these features as *related decorative pictures* and investigate the hypothesis that related decorative pictures affect learning even if they are neither meant to explain text contents nor to distract attention away from the main ideas of the text.

We consider two potential effects of related decorative pictures. First, pictures unspecifically related to a text can have a direct positive effect on learning because they are related to the text's content—without being explanatory or seductive. Second, based on the same idea, related pictures could affect metacognitive monitoring (Metcalf & Shimamura, 1994). These two potential effects will be described in the following sections.

## Effects of related pictures on knowledge acquisition

The pure semantic or associative relatedness of pictures to the learning content can affect learning from text (Bartsch & Cobern, 2003, Experiment 2; Schneider et al., 2018). Bartsch and Cobern (2003) found that pictures related to the text led to better learning than unrelated pictures. Similarly, Schneider et al. (2018) observed in three experiments better knowledge acquisition when participants read text with related instead of unrelated pictures and partially better transfer in conditions with related pictures. Beneficial effects of (positively valenced) related pictures on knowledge and transfer performance were partially also observed in comparisons with an additional condition without any pictures. How can these effects be explained?

In our perspective, there is closely related basic research in cognitive psychology that can parsimoniously account for these findings, although, to our knowledge, these lines of research have hardly referred to each other so far. Davelaar and colleagues provide a parsimonious explanation for better memory performance for related materials compared to unrelated materials based on spreading activation processes (Davelaar et al., 2006). Spreading activation theories (e.g., Anderson, 1976, 1983, 1993) imply that, in semantic networks, activation can spread back and forth between related concepts. Thereby, related concepts mutually reactivate each other (Kowialiewski et al., 2021; Scherer & Wentura, 2018; Schmitz et al., 2014; Schmitz & Wentura, 2012; see also Rohr & Wentura, 2022). This leads to an automatic maintenance of related information in working memory and better performance in working memory tasks (Davelaar et al., 2005, 2006; Scherer & Wentura, 2018, Experiment 2). This automatically extended maintenance of related information in working memory, in turn, enhances the probability that the information is encoded into long-term memory. For unrelated concepts represented in different parts of the semantic

network, there is no multi-directional spread of activation between concepts. Accordingly, this information is less likely encoded into long-term memory (Davelaar et al., 2006). This can explain the more effective knowledge acquisition for text plus related pictures compared to text plus unrelated pictures in the experiments by Bartsch and Cobern (2003) and Schneider et al. (2018). This explanation does not require any explanatory power of the pictures used. Unspecific semantic text-picture-relations are considered sufficient to initiate the postulated processes. To the best of our knowledge, this kind of explanation was yet not referred to in the context of learning from text and pictures.

The effects of a passive mutual facilitation (Davelaar et al., 2006) should primarily be observed with retention measures using knowledge items. Better performance for related materials should not necessarily manifest in transfer or problem-solving performance because of the passive nature of the spreading activation process. If, however, related materials enhance the connectivity among knowledge elements and if learners use their knowledge efficiently, also transfer performance might benefit. However, this is neither a necessary nor a direct effect of a passive mutual activation.

One might be tempted to identify spreading activation effects with the dual coding effect (Clark & Paivio, 1991; Paivio, 1991). Simple dual coding means that the *same* concept is presented visually and verbally and the internal dual coding in a verbal and a non-verbal subsystem is responsible for increased memory performance. However, related decorative pictures do not depict the same concepts as represented in the text. Depicting an open wardrobe, to refer to our introductory example, is not the same as describing how to sort laundry before using a washing machine in the text. Picture and text represent different though related referents. Already Levie and Lentz (1982) stated in their review on effects of text illustrations that "... it appears that the information overlap between a text passage and a facilitative illustration can be something other than 'simple redundancy'" (p. 206). In this perspective, the spreading activation idea extends and specifies to original dual coding hypothesis. We assume that spreading activation processes can explain why non-redundant pictures that only refer to concepts in a text without directly representing these concepts can cause better learning from a text.

## Effects of related pictures on metacognitive monitoring

We assume that related decorative pictures might also increase metacognitive monitoring accuracy. Often, improvements or impairments in memory performance go hand in hand with corresponding effects on metacognitive monitoring (e.g., Barenberg & Dutke, 2019; Souchay, 2007). Therefore, beneficial effects of text-picture relatedness due to spreading activation processes that result in a better encoding into long-term memory could also cause more accurate metacognitive monitoring. This is especially plausible for knowledge items.

This effect should be more likely for metacognitive monitoring accuracy measured in delayed judgments of learning at test (compared to judgments immediately after learning). The reason is: at test, judgments of confidence in answers are based on actual attempts to retrieve information from long-term memory rather than on the expected success of future retrieval (Dunlosky & Nelson, 1992; Little & McDaniel, 2015). Therefore, we used confidence ratings at test as indicators of metacognitive monitoring rather than prospective judgments immediately after learning.

Generally, success and failure of retrieval attempts can be assessed more precisely if representations have been more successfully encoded into long-term memory (e.g., Barenberg & Dutke, 2019). In combination with the assumption that relatedness of information leads to better encoding into long-term memory (Davelaar et al., 2006), text-picture relatedness could increase metacognitive monitoring accuracy measured at test. Summarized, adding related decorative pictures to learning materials could increase performance in knowledge measures and metacognitive monitoring accuracy.

That related pictures might affect metacognitive monitoring is important for at least two reasons. In a theoretical perspective, this might increase the probability of detecting effects of related pictures. Higher confidence in correct responses and lower confidence in false responses, can be a more sensitive measure of learning than measuring for example only the correctness of a response. For example, after retrieval practice, Barenberg and Dutke (2021) found no effects on the number of correct responses, but ratings of confidence in the correctness of answers were higher and less biased. Thus, measuring metacognitive monitoring accuracy can reveal effects of learning that remain undetected otherwise. In an applied perspective, increased metacognitive monitoring accuracy can lead to more efficient regulation of study behavior (e.g., Hines et al., 2009), which can result in better learning (e.g., Mihalca & Mengelkamp, 2020; Thiede et al., 2003).

## Controlling for effects of interest

Beyond potential spreading activation effects, focused in this study, related decorative pictures could also affect learning from text by influencing interest. Two different effects are possible. First, if a decorative picture raises interest only in the depiction itself (but not in the learning contents), attention might be distracted from the text's contents—possibly a case of a “seductive illustration” (see for example, Harp & Mayer, 1997, Experiment 2). Second, if related decorative pictures trigger situational interest (Hidi & Renninger, 2006) not only in the pictures itself, but also in text contents, decorative pictures may improve learning by directing learning activities more efficiently than in situations with lower interest in the learning contents (e.g., Magner et al., 2014).

Therefore, testing the hypothesis that multi-directional spreading of activation in the semantic network can explain facilitative effects of related decorative pictures on learning from text requires controlling for effects of interest. For this purpose, pictures should be used that could be expected to neither increase triggered nor maintained situational interest in the learning contents and triggered and situational interest should be measured as additional dependent variables.

## The current experiments

Increasing evidence suggests that related decorative pictures have different effects on learning from text compared to seductive illustrations and explanatory illustrations. For related decorative pictures, a positive effect on knowledge acquisition (and possibly transfer) might arise that is caused by spreading activation processes and a resulting mutual facilitation of related concepts (Davelaar et al., 2006; Kowialiewski et al., 2021; Scherer & Wentura, 2018). Therefore, we expect a higher percentage of correct responses to knowledge items if participants learn from text with related decorative pictures compared to a condition with

the same text but without pictures (Hypothesis 1). A similar effect is expected for metacognitive monitoring, as mutual facilitation of related concepts based on passive spreading of activation might improve awareness of the availability of learned contents. Therefore, we expect beneficial effects of related decorative pictures on measures of metacognitive monitoring obtained at test compared to the condition without pictures (Hypothesis 2). To test these hypotheses, it is necessary to avoid using explanatory pictures. Further, pictures should not *directly* represent the same main concepts described in a text. This can be achieved straightforwardly if a text describes abstract concepts that cannot be represented directly such as willpower (Experiment 1) or population growth (Experiment 2).

For explorative reasons and for comparability to earlier studies we also measured transfer performance. Given that Hypothesis 1 and/or Hypothesis 2 are corroborated, a positive effect of related decorative pictures on transfer measures would be possible, although transfer depends on more than only the availability of domain knowledge. However, measuring transfer allows us to check for patterns of results incompatible with our hypothesis that beneficial effects of related decorative pictures arise from passive spreading activation processes. A pattern compatible with our hypothesis is (a) both enhanced performance (percentage of correct responses or metacognitive performance) for knowledge items and enhanced transfer performance (percentage of correct responses or metacognitive performance) in the condition with related decorative pictures. Equally compatible is (b) a positive effect (see above) on knowledge but not on transfer. An incompatible pattern would be (c) a positive effect on transfer performance but not on knowledge. This would suggest that the pictures specifically enhance comprehension similar as explanatory pictures do. This would not correspond to our assumption of passive mutual facilitation among related concepts. Another incompatible pattern is given (d) when neither positive effects on transfer nor on knowledge items arise. In addition, all patterns of results with negative effects on knowledge or transfer are incompatible with our assumptions as well.

Further, as a control strategy, we investigated potential effects of related decorative pictures on triggered situational interest (measured during reading) and maintained situational interest (measured after reading). Both measures assessed interest in the text and the topic but not interest in the pictures. Evidence of increased learning performance at the cognitive (Hypothesis 1) and metacognitive level (Hypothesis 2) as a function of decorative pictures could be interpreted more clearly as an effect of passive mutual spreading of activation when *no* positive effects on situational and maintained interest are observed.

## Experiment 1

Experiment 1 investigated the effect of related decorative pictures on learning and metacognitive monitoring in teacher education seminars. Participants read a text about self-control accompanied or not accompanied by related decorative pictures.

## Methods

### Participants and design

In total, data of 79 participants were collected. Performance of eight participants was unexpectedly low (lower than chance level). Their data were excluded from the analyses as it was unlikely that they seriously followed the instructions. Finally, the data of  $n = 36$



**Fig. 1** Examples for related decorative pictures used in Experiment 1. Left side: Photo by bantersnaps / <https://unsplash.com/photos/9o8YdYGT64> / Unsplash Liscence; right side: Photo by Pixabay / <https://www.pexels.com/de/foto/abenteuer-athlet-ausdauer-berge-235922/> / CC0

participants in the condition with related decorative pictures and  $n = 35$  in the control condition were included in the analyses (53 female, 18 male, mean age = 23.9 years,  $SD = 2.8$ ). We conducted the experiment in five teacher education courses. Students were randomly assigned to one of the two conditions: They either read an article on self-control with or without decorative pictures. Interest was measured before, during and after reading. Before reading, prior knowledge was assessed. After reading, knowledge and transfer items were answered. Based on confidence-weighted true–false items (Barenberg & Dutke, 2019) indices of the accuracy of metacognitive monitoring were calculated.

A power analysis using G\*Power (Faul et al., 2007) shows that, in a two-group design, assuming a medium to large effect size of  $d = 0.65$  and  $\alpha = 0.5$  (two-tailed),  $n = 30$  participants per group are needed to achieve a power of  $1 - \beta = .80$ . We estimated this effect size both based on studies with roughly similar methods as well as the cited basic cognitive research with dissimilar methods that, however, investigated the same underlying mechanisms.

## Learning materials

All materials were pen-and-paper-based. The text focused on the resource model of self-control (Baumeister et al., 1998, 2000). It was a shortened version of an article (Baumeister, 2018) from the German science magazine *Spektrum Kompakt*. In the condition with pictures, one to two pictures per page were embedded in the text (11 pictures in total). The text had between 279 and 347 words per page ( $M = 309$ , in total 2472 words on 8 pages).

All pictures used were photos that were thematically associated with the text contents. None of them explained text contents explicitly. Besides the relatedness of the pictures, they decorated the text with primarily aesthetically appealing subjects (Lenzner et al., 2013). For example, one picture showed a library (Fig. 1A), a location closely related to study behavior, which usually requires self-control. Another picture depicted a runner in a mountainous landscape (Fig. 1B). This picture is related to the assumption that self-control requires effort, without explaining precise assumptions, research, or findings addressed in the text or by the following knowledge and transfer items. Another picture showed railroad tracks leading into the distance (related to the future perspective of self-control), healthy food, a person raising their fist into the sky, a bicyclist, smoke (associated with cigarette

consumption and therefore with low self-control), balanced stones piled on top of each other, and a person doing a yoga exercise.

## Measures

Four questions measured *prior knowledge* before reading. Two independent raters evaluated each answer with 0–2 points so that a maximum of 8 points could be achieved. All questions referred to specialist vocabulary or researcher names, for which it is difficult or impossible to provide meaningful answers if one has never obtained scientific knowledge about self-control. The raters agreed on 93–100 percent of their ratings. Therefore, their ratings were averaged.

Also before reading, *topic interest* was measured using 7 items (Cronbach's  $\alpha = .84$ ) from the extended version of a questionnaire by Schaffner and Schiefele (2008, modified by Schiefele, personal communication, August 3, 2018). These items target the feeling and value components of topic interest. Responses were given on a four-point Likert scale. The mean rating was used as a measure of topic interest. Before answering, participants were informed that the text will deal with willpower, also known as self-control or self-regulation and that all these terms refer to our ability to work persistently towards one's own goals. They were informed that the items described different attitudes they can have on this topic.

During reading, *triggered situational interest* was measured with a single item ("... please briefly state how interesting you find the section of the article you have just read") presented at the bottom of each page (Cronbach's  $\alpha = .91$ ). At the end of each page, participants were instructed to stop and indicate how interesting they found the text content they have just read. A 7-point response scale with ratings from "not at all interesting" to „very interesting“ was used. The 8 ratings per participant (one rating per page) were averaged to obtain a global measure of situational interest.

*Maintained situational interest* was measured after reading. Participants were informed that the experimenter wanted to know about which text topics they liked to learn more. They were asked to imagine that they would attend another seminar in the same module. Then, they indicated for each of six described topics<sup>1</sup> whether it should be included in the future seminar. For each topic, participants indicated "yes" (it should be added to the planned seminar) or "no" (they are not interested in it).

As an additional measure, an *open comprehension question* was applied. For this question, first, a case description (5 sentences) described student behavior. Then, participants were asked to explain the behavior based on the resource model of self-control. The quality of the answers was rated (same procedure as for prior knowledge items). Raters agreed in

<sup>1</sup> The six topics were described to the students in some detail. Here we name the topics shortly. The first was described as an alternative theory to explain many of the findings described in the experimental text. The second was described as a passage from a book, which is not directed at specialist readers similar to the article read, from the author of the article. The third was an article from the same science magazine as the text the students just read and included an interview with Walter Mischel whose "Marshmallow-Experiment" was also mentioned in the experimental text. The fourth topic was a study about the relevance of self-control. The fifth topic was a theory which can be regarded as a kind of overarching model of self-control (additionally it was stated, that from this theory some implications for interventions strengthening self-control can be derived). The sixth topic was a simple technique, to help follow your intentions in situations where self-control is required.

85.92 percent of their ratings and the ratings were highly correlated ( $r = .94, p < .001, 95\%$  CI [.88, .98]).<sup>2</sup>

Participants answered 11 knowledge and 11 transfer items. Items were randomly ordered with the same order for all participants. We used confidence-weighted true–false items (Dutke & Barenberg, 2015) with four choices. In these items, participants indicate (1) that they are sure, that the presented statement is incorrect, (2) that they think that the statement is incorrect, but that they are unsure, (3) that they think that the statement is correct, but they are unsure, or (4) that they are sure that the statement is correct. Thereby, the answer provided information about its correctness and the participant’s confidence in the correctness of this response. Whereas knowledge items referred to a particular part of the text, transfer items required the reader to establish relations between different parts of the text or, alternatively, an integration of the propositions given in the text with prior knowledge (cf. Jaeger & Wiley, 2014; Wiley et al., 2005).

To measure performance, the percentage of correct responses (irrespective of the confidence ratings) was computed separately for knowledge and transfer items. To measure metacognitive monitoring, three indices were computed separately for responses to knowledge and transfer items based on formulas (see appendix) provided by Barenberg and Dutke (2019). The *absolute accuracy* (AC) of confidence judgments was calculated by adding the proportion of correct and confident responses to the proportion of incorrect and unconfident responses. Thus, AC measures the match of correctness and confidence. This index reflects the precision of the confidence judgments across all items of one type (either knowledge or transfer). Second, we analyzed how reliable participants discriminated between correctly and incorrectly answered items on the basis of their confidence ratings. *Discrimination* (DIS) was calculated by subtracting the unwarranted confidence in the correctness of actually incorrectly answered items from the justified confidence in correctly answered items. Confidence in correctly answered items was weighted by the number of actually correctly answered items, and confidence in incorrectly answered items was weighted by the number of incorrectly answered items, before the difference was calculated. High DIS scores indicate successful metacognitive monitoring. A DIS score of zero would indicate equal confidence in answers that are actually correct and in answers that are actually incorrect.

Assuming better metacognitive monitoring in the condition with related decorative pictures, both the absolute accuracy (AC) and the discrimination score (DIS) should be higher in the condition with related decorative pictures compared to the control condition. These comparisons presuppose that both groups over- or underestimate their learning performance to a similar degree. Therefore, we calculated a third index of metacognitive monitoring, *bias* (BS), as an additional variable by subtracting the relative number of correctly answered items from the relative number of confidently answered items. This difference is negative, when more items were correctly answered than items were answered with confidence (indicating underconfidence). Correspondingly, a positive difference indicates overconfidence.

<sup>2</sup> In parentheses: the 95% confidence interval from the bootstrapping procedure (with 1000 bootstrap samples).



## Procedure

We conducted the experiment at university during 90-min seminar sessions. Students were informed that participating was voluntary and that all data will be analyzed anonymously. Independent experimenters managed the data collection. Seminar lecturers were absent during the study.

Participants within each seminar were randomly assign to one of two groups working in different rooms. During answering the items and text reading, participants were not allowed to turn back to already read pages. First, participants' demographic data were assessed, followed by the measure of prior knowledge and the topic interest questionnaire. Then, participants read the text. They were instructed that they can read for a maximum of 20 min. The experimenter ensured that students had the opportunity to read the complete text, even when they need more than 20 min.<sup>3</sup> Triggered situational interest was measured eight times during reading. After reading, participants first completed the maintained situational interest measure. Next, they answered the open comprehension question, followed by knowledge and transfer items. Afterward, participants gave written consent that their anonymized data will be used for scientific purposes.

## Results

### Pre-analyses

A Mann–Whitney-U-Test did not reveal a significant difference in prior knowledge between the group with ( $Mdn=0$ ) and without related decorative pictures ( $Mdn=0$ ),  $U(35,36)=645.00$ ,  $p=.845$ . Prior knowledge significantly correlated with the proportion of correct responses to knowledge items ( $r=.33$ ,  $p=.005$ , 95% CI [.17, .49]) but not significantly with the proportion of correct responses to transfer items ( $r=.23$ ,  $p=.050$ , 95% CI [-.04, .42]).

Topic interest did not differ between the condition with ( $M=2.02$ ,  $SD=0.51$ ) and without related decorative pictures ( $M=1.97$ ,  $SD=0.48$ ),  $t(69)=0.43$ ,  $p=.665$ . However, topic interest correlated significantly with both triggered situational interest ( $r=.36$ ,  $p=.002$ , 95% CI=[.16, .54]) and maintained situational interest ( $r=.24$ ,  $p=.042$ , 95% CI=[.06, .43]).

### Main Analyses

To analyze the effects of related decorative pictures on measures of memory performance, prior knowledge was included as a potential covariate. Analyzing prior knowledge revealed that the covariate was not normally distributed. Because most participants (60.6%) obtained 0 points on the prior knowledge measure and the rest obtained only few points, this variable mainly differentiated between participants having some prior knowledge of the learning subject ( $>0$  points) and participants having no knowledge ( $=0$  points). Therefore, we used prior knowledge as a dichotomized variable (with and without prior knowledge).

*Open comprehension question.* The performance in the open comprehension question was analyzed in an ANCOVA with the factor presence vs. absence of related decorative

<sup>3</sup> However, for the first seminar tested (11 participants) indeed only 20 min were given for text reading.

**Table 1** Effect on the metacognitive measures calculated for knowledge items (Experiment 1)

Measure	<i>M</i> ( <i>SD</i> ) without decorative pictures	<i>M</i> ( <i>SD</i> ) with decorative pictures	<i>F</i>	<i>p</i>	$\eta_p^2$
AC	.68 (.14)	.69 (.16)	$F(1, 68) = .24$ [1.96]	.628 [.166]	.003 [.028]
DIS <sup>a</sup>	.41 (.21)	.46 (.22)	$F(1, 58) = 1.10$ [3.84]	.298 [.055]	.019 [.019]

In square brackets the corresponding values for the effect of the covariate (prior knowledge) are indicated

<sup>a</sup> 7 Participants in the condition with related decorative pictures, and 3 participants in the condition without pictures did not give incorrect answers to knowledge items. Therefore, the discrimination score (DIS) could not be computed for those participants

pictures and the covariate presence vs. absence of prior knowledge. Participants in the group with decorative pictures gave on average  $M = 58.68$  ( $SD = 43.80$ ) percent of correct answers compared to  $M = 52.14$  ( $SD = 46.13$ ) percent of correct answers in the control condition with no significant main effect for the factor presence of related decorative pictures,  $F(1, 68) = 0.32$ ,  $p = .572$ ,  $\eta_p^2 = .005$ . The effect of the covariate prior knowledge was also not significant,  $F(1, 68) = 0.34$ ,  $p = .564$ ,  $\eta_p^2 = .005$ .

**Knowledge.** Performance for knowledge items was calculated as the percentage of correct answers on knowledge items (irrespective of the confidence in the correctness of these answers). An ANCOVA with the factor presence vs. absence of related decorative pictures and prior knowledge as a covariate revealed that participants in the condition with related decorative pictures gave significantly more correct answers ( $M = 83.40$ ,  $SD = 11.90$ ) compared to the control condition ( $M = 77.40$ ,  $SD = 13.62$ ),  $F(1, 68) = 5.51$ ,  $p = .022$ ,  $\eta_p^2 = .075$ . The covariate prior knowledge also reached significance with higher knowledge scores for participants with ( $M = 86.05$ ,  $SD = 9.19$ ) than for participants without prior knowledge ( $M = 76.79$ ,  $SD = 13.95$ ),  $F(1, 68) = 11.25$ ,  $p = .001$ ,  $\eta_p^2 = .142$ . As dichotomizing the covariate reduces its variance, we conducted an additional analysis with prior knowledge as a continuous covariate. In this ANCOVA, the effect of the factor presence vs. absence of related decorative pictures was significant,  $F(1, 68) = 3.77$ ,  $p = .028$  (one-tailed),  $\eta_p^2 = .053$  ( $F(1, 68) = 8.22$ ,  $p = .006$ ,  $\eta_p^2 = .108$ , for the effect of the covariate).

**Transfer.** The percentage of correct answers on transfer items (irrespective of the confidence) was analyzed in an ANCOVA with the factor presence vs. absence of related decorative pictures and prior knowledge as a covariate. The analysis revealed no significant difference between the condition with ( $M = 67.45$ ,  $SD = 16.44$ ) and without related decorative pictures ( $M = 67.25$ ,  $SD = 15.27$ ),  $F(1, 68) < 0.01$ ,  $p = .945$ ,  $\eta_p^2 < .001$ . There was no effect of the covariate,  $F(1, 68) = 0.05$ ,  $p = .826$ ,  $\eta_p^2 = .001$ .

**Metacognitive monitoring of knowledge items** (Table 1). Both groups of students slightly underestimated the correctness of their answers, but the bias (BS) did not differ significantly between the groups learning with ( $M = -.10$ ,  $SD = .11$ ) and without ( $M = -.07$ ,  $SD = .21$ ) pictures,  $F(1, 68) = 0.74$ ,  $p = .393$ ,  $\eta_p^2 = .011$ . Thus, the metacognitive variables AC and DIS can be directly compared between the groups. For both measures, ANCOVAs with the factor presence of the pictures and the covariate presence of prior knowledge were performed, showing no significant effects.

**Metacognitive monitoring of transfer items** (Table 2). Both groups slightly underestimated the correctness of their answers, but the bias (BS) did not differ between the groups

**Table 2** Effect on the metacognitive measures calculated for transfer items (Experiment 1)

Measure	<i>M</i> ( <i>SD</i> ) without decorative pictures	<i>M</i> ( <i>SD</i> ) with decorative pictures	<i>F</i>	<i>p</i>	$\eta_p^2$
AC	.57 (.16)	.57 (.19)	$F(1,68)=0.00$ [.94]	.979 [.336]	.000 [.014]
DIS <sup>a</sup>	.23 (.24)	.24 (.28)	$F(1, 66)=0.40$ [.33]	.842 [.579]	.001 [.005]

In square brackets the corresponding values for the effect of the covariate (presence of prior knowledge) are indicated

<sup>a</sup> One participant in the condition with related decorative pictures and one participant in the condition without pictures did not give incorrect answers to transfer items. Therefore, the discrimination score (DIS) could not be computed for these participants

learning with ( $M = -.09$ ,  $SD = .23$ ) and without ( $M = -.11$ ,  $SD = .25$ ) related decorative pictures,  $F(1,68) = 0.12$ ,  $p = .735$ ,  $\eta_p^2 = .002$ . Thus, the variables AC and DIS can be directly compared between both groups. For both measures ANCOVAs with the factor presence of related decorative pictures and the covariate prior knowledge were performed, showing neither significant effects of the pictures nor an effect of the covariate.

**Interest.** For triggered and maintained situational interest measures ANCOVAs with the factor presence of pictures and topic interest as a covariate were performed.

For triggered situational interest no significant difference between the condition with ( $M = 5.57$ ,  $SD = 0.76$ ) and without pictures ( $M = 5.65$ ,  $SD = 0.93$ ) was observed,  $F(1, 68) = 0.05$ ,  $p = .815$ ,  $\eta_p^2 = .001$ . As the significant positive correlation in the pre-analyses suggests, the covariate had a significant positive effect,  $F(1, 68) = 9.84$ ,  $p = .003$ ,  $\eta_p^2 = .126$ .

To ensure, that triggered situational interest was not influenced by the pictures, we additionally performed the corresponding Bayesian analysis of covariance on triggered situational interest, including pictures (with vs. without) as a fixed factor and topic interest as a covariate. The Bayesian ANCOVA works by comparing four models with varying predictors of maintained situational interest: (1) a null model; (2) a model containing only pictures as a predictor; (3) a model containing only topic interest as a predictor and (4) a model containing both pictures and topic interest as predictors. Only Model 3, predicting triggered situational interest only by topic interest measured before reading, had its model odd increased after observing data ( $BF_M = 8.14$ ). Therefore, Model 3 was most probable,  $P(M|data) = .73$ . and the observed data was 3.47 times more likely under Model 3 than under the next probable model, which was Model 4.

To account for model uncertainty, we performed Bayesian model averaging to test the “effects” of both predictors. The data were 15.89 times more likely under models containing topic interest as a predictor, but only 0.29 times as likely when including the factor presence of pictures. Thus, we concluded that only topic interest prior to reading was related to triggered situational interest. Critically, the presence of the pictures had no effect.

For maintained situational interest (the amount of related seminar materials chosen for a future seminar), no significant difference between the condition with ( $M = 4.50$ ,  $SD = 1.08$ ) and without pictures ( $M = 4.60$ ,  $SD = 0.88$ ) was found,  $F(1, 68) = 0.11$ ,  $p = .743$ ,  $\eta_p^2 = .002$ . The covariate had a significant positive effect on maintained situational interest,  $F(1, 68) = 4.16$ ,  $p = .045$ ,  $\eta_p^2 = .058$ .

To ensure, that maintained interest is not influenced by the pictures in this experiment, we performed additionally a corresponding Bayesian analysis of covariance on maintained situational interest including pictures (with vs. without) as a fixed factor and the covariate topic interest. The same four models with varying predictors of maintained situational interest were compared as in the analysis of triggered situational interest. The model containing only topic interest as a predictor (3) and the null model (1) had their model odds increased after observing data ( $BF_M=2.74$  and  $BF_M=1.39$ , respectively). Of these, Model 3 was most probable,  $P(M|data)=.48$ , and the observed data was 1.50 times more likely under Model 3 than under Model 1.

To account for model uncertainty, we performed Bayesian model averaging to test the “effects” of both predictors. The data were 1.49 times more likely under models containing topic interest as a predictor, but only 0.26 times as likely when including the factor presence of pictures. Thus, we concluded that only topic interest was related to maintained interest. Critically, the presence of the pictures had no positive effect on maintained situational interest.

Please note that the clear absence of picture effects on interest provides a solid hint that any mediating effects of interest are highly unlikely. In many approaches, for example, the causal steps approach by Baron and Kenny (1986), the mere insignificance of these effects was regarded sufficient to rule out mediation. As this approach has recently been criticized, we provide additional indirect tests of mediation using PROCESS v4.1 for SPSS (Hayes, 2017) in Supplement S1 (Online Resource 1) to test for the dependent variable for which we have obtained an effect of related pictures (the percentage of correct responses to knowledge items), whether interest mediated this effect. This analysis did not provide any indications of a mediation through interest.

Additionally, we analyzed the correlation of maintained situational interest with the other interest measures. Maintained situational interest, operationalized as the number of seminar materials related to the text which participants would choose for a fictitious seminar, correlated significantly with both topic interest ( $r=.24$ ,  $p=.042$ , 95% CI=[.06, .43]), and triggered situational interest ( $r=.33$   $p=.005$ , 95% CI=[.10, .54]).

## Discussion

Experiment 1 shows a positive effect of related decorative pictures on memory performance for knowledge items (Hypothesis 1), but no effect on metacognitive measures (Hypothesis 2) and transfer performance. Presence of the used pictures did neither increase triggered nor maintained situational interest and interest did not mediate the picture effect on knowledge acquisition.

The main finding is a beneficial effect on memory performance by pictures related to the topic of the text. This can be explained by a passive mutual facilitation of the concepts of related materials in working memory that transits to long-term memory (Davelaar et al., 2006).

The used pictures did neither increase triggered situational interest nor maintained situational interest. Therefore, their main influence is likely on memory and not on interest. This pattern may differ when other pictures are presented, but in the present experiment interest cannot account for better memory in the pictures condition.

We used a variant of an ecologically valid measure of maintained situational interest. Whereas many studies use simple verbal measures (e.g., Harp & Mayer, 1997), our participants indicated for multiple topics whether these topics should be part of another seminar

they could participate in. This measure correlated with a more established measure of topic interest and with triggered situational interest measured during reading. Thus, this behavioral measure of maintained situational interest seems promising for future research.

In conclusion, Experiment 1 implies that non-explanatory decorative pictures that are thematically related to the text promote learning. This effect is probably not mediated by higher interest. Instead, the most parsimonious explanation is a mutual facilitation of related concepts based on spreading activation processes that lead to deeper encoding of related content into long-term memory.

## Experiment 2

### Methods

In Experiment 2, we essentially employed the same procedure and pursued the same hypotheses as in Experiment 1. As Experiment 2 was planned before the data collection for Experiment 1 was completed, the same effect size as in Experiment 1 was assumed ( $d=0.65$ ). We conducted this experiment with a younger sample and modified learning materials to assess the robustness and generalizability of our results. The main changes in the procedure were that, due to the younger sample, we used a different text with a different topic and less knowledge and transfer items. A different but conceptually similar measure for maintained interest was applied. In addition, there was no open comprehension question as applied in the first experiment.

### Participants and design

In total, data of 75 eight-graders from three secondary school classes were collected. All these students and their parents consented that the data are used anonymously for scientific purposes. The data of participants ( $n=19$ ) who did not perform above chance level were excluded from the analyses resulting in  $n=29$  (with pictures) and  $n=27$  (without pictures). The mean age of the final sample (25 female, 31 male) was  $M=13.62$  ( $SD=0.65$ ). Students were randomly assigned to the two conditions.

Again, a one-factorial design was used with presence vs. absence of related decorative pictures as a between-participants factor. Prior knowledge and (prior) topic interest were measured as potential covariates. Knowledge acquisition, transfer, and metacognitive monitoring were measured as dependent variables and the same forms of interest as in Experiment 1 were measured as additional dependent variables (to rule out alternative explanations).

### Learning material

All materials were pen-and-paper-based. The text consisted of 5 pages (563 words) dealing with different facets of population growth and related topics. Each page had 98 to 129 words ( $M=112.6$  words per page). The text contained passages from schoolbooks and a didactic journal (Bethke, 2009; Bremm & Latz, 2012; Frank et al., 2009) updated with more recent data (Statista, 2018a, 2018b).

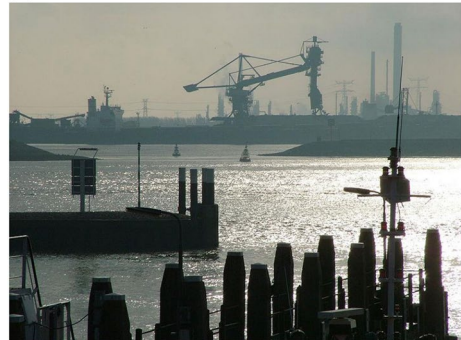
Again, all related decorative pictures used were thematically associated with the learning material without explicitly explaining text contents. For example, a satellite picture of

**Fig. 2** Examples for a related decorative picture used in Experiment 2. A: WikiImages / <https://pixabay.com/de/photos/erde-blauer-planet-globus-plane-ten-11015/> / Pixabay Licence; B: Ziko van Dijk / [https://commons.wikimedia.org/wiki/File:Industrie\\_in\\_Hoek\\_van\\_Holland,\\_2005.jpg](https://commons.wikimedia.org/wiki/File:Industrie_in_Hoek_van_Holland,_2005.jpg) / CC BY-SA 3.0

(A)



(B)



the earth was used that was loosely related to the topic of global population growth (see Fig. 2A). This picture does not explain anything about global population growth, for example it does not highlight developed or less developed countries. Similarly, an illustration depicting industrial facilities decorated a paragraph about low birthrates in industrialized states (see Fig. 2B). Another picture showed a beach with people, another a water tap, another bananas (related to developing countries).

## Measures

Three open questions measured *prior knowledge*. These questions addressed specialist vocabulary (such as "carrying capacity"), for which it is unlikely that students know them if they have never learned anything about the subject. As in Experiment 1, two independent raters awarded 0–2 points for each answer. As they agreed on 91.10–100% of their ratings, these were averaged across raters.

The same measure for *topic interest* as in Experiment 1 was used (Cronbach's  $\alpha$ , in this sample = .92). Again, before participants rated interest in the topic of the text, the topic was briefly described. Participants were informed that the text addressed the development of the population on earth, that is the development of the number of people on earth that is influenced, for example, by birth and death rates as well as immigration and emigration.

Participants rated their *triggered situational interest* after reading each page (Cronbach's  $\alpha = .85$ ) on a 7-point response scale from "not at all interesting" to "very interesting."

We used *confidence-rated true–false items* to measure memory performance with  $N=10$  knowledge items presented first, followed by  $N=10$  transfer items. The same metacognitive indices were calculated for knowledge and transfer items as in Experiment 1.

*Maintained situational interest* was measured in a similar choice situation as in Experiment 1. Participants were informed that in front of the classroom there were six different short texts that they could take with them. Then, the topics of these texts were described in one sentence. The description made clear that all texts are related to the topic of the text read before. The additional short take-home texts covered population policy in China (from Claaßen, 2016), educational efforts in Tanzania (taken from Hoppe et al., 2018), the carrying capacity of the earth (from Bethke, 2009), family policy in Germany (from Bösch, 2012), world food (from Schweighöfer & Herre, 2017) and demographic change in Japan (from Claaßen, 2015). Participants indicated with yes or no for each text whether they wanted to take it with them. They were not able to inspect the texts before making their decision. They were instructed that they could take as many texts as they liked. The measure of maintained situational interest was the number of yes-responses per participant.

## Procedure

The experiment was conducted in classrooms during 90-min sessions. Within each class, participants were randomly assigned to the experimental or the control condition. Participants in both groups were tested in the same room. Screening walls were placed between participants.

Prior knowledge and topic interest were measured before reading. Then, participants received the text. It was stated that they could read for up to 35 min. The experimenter ensured that students had the opportunity to read the complete text, even if more time was needed. After text reading and indicating triggered situational interest, participants answered the knowledge and transfer items. Then, students' maintained situational interest was measured, before they gave written consent that their data will be used for scientific purposes.

## Results

### Preanalyses

A Mann–Whitney-U-Test did not reveal a significant difference in prior knowledge between the group with ( $Mdn=0.00$ ) and the group without pictures ( $Mdn=0.00$ ),  $U=304.50$ ,  $p=.095$ . Further, prior knowledge did neither significantly correlate with the percentage of correct responses to knowledge items ( $r=.01$ ,  $p=.935$ , 95% CI=[-.25, .34]) nor with the percentage of correct responses to transfer items ( $r=.08$ ,  $p=.539$ , 95% CI=[-.16, .32]).

A significant difference in topic interest measured before participants read the text was observed, with higher topic interest in the condition with pictures ( $M=2.57$ ,  $SD=0.70$ ) than in the condition without pictures ( $M=2.10$ ,  $SD=0.64$ ),  $t(54)=2.65$ ,  $p=.011$ . Topic interest correlated significantly with triggered situational interest ( $r=0.72$ ,  $p<.001$ , 95% CI=[.57, .82]) and maintained situational interest ( $r=.39$ ,  $p=.003$ , 95% CI=[.14, .61]).

**Table 3** Effect on the metacognitive measures calculated for knowledge items (Experiment 2)

Measure	<i>M (SD)</i> without decorative pictures	<i>M (SD)</i> with decorative pictures	<i>t</i>	<i>p</i>
AC	.57 (.17)	.70 (.20)	$t(54)=2.56$	.013*
DIS	.29 (.22)	.44 (.25)	$t(54)=2.31$	.025*

\*  $p < .05$ **Table 4** Effect on the metacognitive measures calculated for transfer items (Experiment 2)

Measure	<i>M (SD)</i> without decorative pictures	<i>M (SD)</i> with decorative pictures	<i>t</i>	<i>p</i>	<i>d</i>
AC	.51 (.15)	.51 (.17)	$t(54)=0.15$	.882	0.04
DIS	.05(.16)	.08 (.18)	$t(54)=0.67$	.503	0.18

## Main analyses

**Knowledge.** As the preanalyses revealed, prior knowledge did neither correlate with performance for knowledge items nor with transfer performance. Therefore, prior knowledge was not used as a covariate.

For performance on knowledge items (percentage of correct responses regardless of the confidence in the correctness of the answers), there was no significant difference between the group with related pictures ( $M=73.79$ ,  $SD=14.25$ ) and the group without pictures ( $M=71.85$ ,  $SD=13.60$ ),  $t(54)=.52$ ,  $p=.605$ ,  $d=0.14$ .

**Transfer.** The corresponding analysis of transfer performance also showed no significant difference in performance between the condition with pictures ( $M=56.55$ ,  $SD=12.61$ ) and the control condition ( $M=54.07$ ,  $SD=13.09$ ),  $t(54)=0.72$ ,  $p=.474$ ,  $d=0.19$ .

**Metacognitive monitoring of knowledge items.** Both groups of students slightly underestimated the correctness of their answers, but BS did not differ between the groups learning with ( $M=-.17$ ,  $SD=.20$ ) and without ( $M=-.15$ ,  $SD=.25$ ) related decorative pictures,  $t(54)=-0.28$ ,  $p=.780$ ,  $d=0.08$ . Thus, AC and DIS can be directly compared between groups: Participants in the condition with related decorative pictures had a higher absolute accuracy (AC) of confidence judgments and showed a better discrimination of correct and incorrect responses (DIS) based on the confidence ratings (see Table 3).

**Metacognitive monitoring of transfer items.** Both groups of students slightly underestimated the correctness of their answers, but BS did not differ between the groups learning with ( $M=-.10$ ,  $SD=.21$ ) and without ( $M=-.09$ ,  $SD=.25$ ) pictures,  $t(54)=.12$ ,  $p=.906$ ,  $d=0.03$ . Thus, AC and DIS can be directly compared between groups. No significant differences regarding the metacognitive measures for transfer items between the two conditions were found (Table 4).

**Interest.** For all following analyses of interest measures, ANCOVAs with the between-participants factor presence of pictures and the covariate topic interest were performed.

For triggered situational interest no significant difference between the condition with ( $M=4.86$ ,  $SD=1.10$ ) and without related decorative pictures ( $M=4.42$ ,  $SD=1.22$ ) was observed,  $F(1, 53)=0.40$ ,  $p=.532$ ,  $\eta_p^2=.007$  ( $F(1, 53)=53.07$ ,  $p=.000$ ,  $\eta_p^2=.500$ , for the effect of the covariate).



To ensure that triggered situational interest is not influenced by the pictures, we additionally performed a corresponding Bayesian analysis of covariance (ANCOVA) on triggered situational interest, including pictures (with vs. without) as a fixed factor and topic interest as a covariate. Four models with varying predictors of triggered situational interest were compared: (1) a null model; (2) a model containing only presence of pictures as a predictor; (3) a model containing topic interest as a predictor and (4) a model containing both pictures and topic interest as predictors. The model with topic interest as the only predictor (3) had its model odd increased after observing data ( $BF_M=9.60$ ). This model was most probable,  $P(M|data)=.76$ ; the observed data was 3.19 times more likely than under the next probable model, which was Model 4.

Accounting for model uncertainty, we performed Bayesian model averaging to test the “effects” of both predictors. The data were 1.187e7 times more likely under models containing topic interest as a predictor, but only 0.31 times as likely when including the factor pictures. We concluded that only topic interest prior to reading is related with triggered interest. Critically, the presence of the used related decorative pictures had no effect.

Regarding maintained situational interest, participants in the group with related decorative pictures indicated to take less materials with them to read ( $M=1.62$ ,  $SD=1.54$ ) compared to the control group ( $M=2.00$ ,  $SD=1.39$ ),  $F(1, 53)=5.56$ ,  $p=.022$ ,  $\eta_p^2=.095$ . The covariate had a significant positive effect,  $F(1, 53)=14.44$ ,  $p<.001$ ,  $\eta_p^2=.214$ . Because of the *negative* effect, there is clearly no significant *positive* effect of the pictures on maintained situational interest: No Bayesian ANCOVA is necessary to confirm this.

As in the first experiment, the clear absence of these effects indicated that any mediating effects of interest are highly unlikely (cf. Baron & Kenny, 1986). Nevertheless, we conducted mediation analyses (Supplement S1, Online Resource 1) for those dependent variables which showed an effect of related pictures (that are AC and DIS). These analyses did not provide any indication of a mediation through interest.

Moreover, we analyzed correlations of maintained situational interest with the other interest measures. As expected, maintained situational interest, the number of short texts that participants wanted to take, correlated significantly with both topic interest ( $r=.39$ ,  $p=.003$ , 95% CI=[.14, .61]) and triggered situational interest ( $r=.39$ ,  $p=.003$ , 95% CI=[.14, .60]).

## Discussion

In Experiment 2, students in school learned from a text about population growth. Non-explanatory, related decorative pictures were added to the text in the experimental condition. Regarding measures of knowledge acquisition and transfer, results showed no effects of related decorative pictures on the percentage of correct responses, neither for knowledge (Hypothesis 1) nor for transfer items. However, related pictures increased the accuracy of metacognitive monitoring for knowledge items (not for transfer items). Participants performed significantly better in the condition with related decorative pictures compared to the control condition regarding both, the overall accuracy of metacognitive judgments (AC) and the discrimination of correctly and incorrectly answered items (DIS) on the basis of the confidence ratings (Hypothesis 2).

Summarized, in Experiment 2, for knowledge items, the accuracy of the confidence judgments benefitted. Again, no effects of the pictures on transfer measures emerged – neither on the correctness of answers nor on the accuracy of confidence judgments.

This pattern of results can be explained by a mutual facilitation of related concepts in working memory that transits to long-term memory (Davelaar et al., 2006).

The related decorative pictures did not increase interest. Instead, they led to *less* maintained situational interest. This is a puzzling finding: Maintained situational interest was reduced by the pictures, although triggered situational interest was not influenced. However, Hidi and Renninger (2006) hypothesized that maintained situational interest arises from triggered situational interest. This could imply that our measure of maintained situational interest has a higher sensitivity than the measure of triggered situational interest or that the effect is an incidental finding. We assume that a central advantage of the maintained situational interest measure we used is that it is not a merely verbal measure of whether something is considered as “interesting” or not, but a direct measure of whether participants want to engage further with this topic. Thus, this measure provides a direct link to future interest-related behavior although it is no explicit self-judgment of interest. Importantly, results of both interest measures imply that interest does not seem to be a candidate to explain the effect of related decorative pictures on metamemory.

## General discussion

In two experiments, one at university (with young adults), one in school (with eight-grade students), participants learned from text decorated with non-explanatory but thematically related pictures (experimental condition) or not decorated (control condition).

The effects on the memory and metacognitive measures are in line with the assumption of a mutual facilitation of related concepts caused by spreading-activation processes (e.g., Davelaar et al., 2006). In both experiments, related pictures had positive effects on measures derived from knowledge items. In Experiment 1, reading text with related decorative pictures led to more correct responses to knowledge items (Hypothesis 1), but not to increased metacognitive monitoring accuracy (Hypothesis 2). In Experiment 2, for knowledge items, positive effects of related decorative pictures on metacognitive monitoring accuracy were observed (Hypothesis 2), but no effect on the number of correct responses to knowledge items (Hypothesis 1). In both experiments, the presence of related pictures had no effect on measures derived from transfer items.

In both experiments, the pictures did not increase interest in the text. Triggered situational interest did not differ between conditions. Maintained situational interest did not differ between conditions in Experiment 1. In Experiment 2, decorative pictures reduced maintained situational interest. This makes it unlikely that the effects of the pictures on memory and on metacognitive monitoring can be explained by increased interest due to the presence of decorative pictures.

## Effects of related pictures on memory and metacognition

The positive effect of related pictures regarding knowledge manifested itself in better memory in Experiment 1 and in better metacognitive monitoring in Experiment 2. Therefore, the findings of both experiments are in line with both basic research on effects of relatedness (Davelaar et al., 2006; Scherer & Wentura, 2018) and applied

research on learning with text and pictures (see Bartsch & Cobern, 2003; Schneider et al., 2018) suggesting that semantic relatedness improves memory and learning outcomes. This convergence is remarkable as, until now, these lines of research were largely unconnected.

The effects on knowledge items are not trivial for two reasons. First, the pictures did not explain basic ideas of the text. Second, we chose abstract and complex topics (self-control and population development) that are difficult to represent in pictures. Consequently, the pictures were only weakly related to the text. One explanation for the positive effects is thus automatic mutual facilitation of the representations of related materials based on spreading activation processes (Davelaar et al., 2006; Scherer & Wentura, 2018). Following this assumption, activation can spread back and forth among related concepts once they are active in working memory. Thus, mutual facilitation of related concepts, in this case text and picture content, is assumed to occur in working memory. This process leads to a longer activation of related concepts compared to unrelated concepts. This, in turn, causes better encoding into long-term memory. Thereby, spreading activation processes can explain the better performance in the condition with related decorative pictures for knowledge items. Under which conditions these activation processes express themselves in higher recognition of learned contents or more adequate confidence in recognized contents is beyond the scope of this research and requires a more thorough analysis of motivational aspects of the test situation and therefore different experimental designs. However, a preliminary interpretation of the difference in results between Experiment 1 und Experiment 2 may refer to a bottom effect as the overall performance in Experiment 2 was comparably low.

The absence of effects on transfer measures in both experiments is compatible with our assumption of a mutual facilitation of related concepts and automatic spreading activation processes. Higher availability of text content knowledge due to spreading of activation may or may not facilitate transfer. We included transfer measures to rule out that related pictures increased transfer without increasing the availability of content knowledge, which would have been incompatible with our theorizing. The pattern of results we found in our experiments (increased content knowledge without increased transfer performance) demonstrates that related decorative pictures probably influence learning in a different way than explanatory pictures (e.g., Mayer, 1989).

Although overall-results are clearly compatible with our hypotheses, the results should not be over-interpreted. The effects should not lead to the conclusion that decorative pictures per se increase learning for all text-picture-combinations. As meta-analyses of the seductive detail effect show, illustrations can have negative effects on learning (Rey, 2012; Sundararajan & Adesope, 2020). There might be a fine line between pictures contributing to mutual activation processes so that learning benefits and pictures attracting so much attention that text learning suffers. A practical implication may be that non-explanatory illustrations, at least when they are semantically related to the text, are not necessarily detrimental to learning.

The pattern of results we found could also be interpreted as an effect of dual coding. A basic assumption in this theoretical context states that the same concept represented in a pictorial and a verbal code can be easier retrieved than a concept represented only in the verbal or the non-verbal subsystem. In our experiments, however, the pictures did not represent the same concepts as represented in the text but only loosely referred to the same semantic field that was also touched by the concepts of the text. Nevertheless, even under this condition, pictures improved learning. We propose that spreading activation processes provide the missing link: The activation of concepts by the pictures support the activation of concepts of the text although these concepts were not identical.

## Interest measures

Here, interest measures served as additional dependent variables to rule out an alternative explanation. As the pictures in the experimental condition did not increase situational interest, it is improbable that increased interest mediated the effect of the pictures on memory. Spreading of activation mechanisms, as hypothesized, are the more probable explanation.

Whereas, many studies measured maintained interest verbally (e.g., Magner et al., 2014; Schiefele, 1990), we used an ecologically more valid measure – the participants' choices (of related materials). Remarkably, Experiment 2 shows a negative effect on maintained situational interest but no effect on triggered situational interest. This pattern of results is incompatible with the four-phase model of interest development (Hidi & Renninger, 2006) according to which a negative effect on maintained situational interest would have required a corresponding effect on the previous phase of interest development. The pattern of results in Experiment 2 might be related to the way we measured maintained interest. Although this measure correlated positively with the verbal interest measures, it might have characteristics (higher sensitivity or situation specificity) that need replication and further exploration.

## Limitations

In the current studies, no condition with unrelated pictures was used. However, unrelated pictures would introduce additional cognitive load and may cause a seductive detail effect (Rey, 2012; Sundararajan & Adesope, 2020). Accordingly, the observed effects would probably be more pronounced when comparing performance in a condition with related pictures with performance in a condition with unrelated pictures. In our view, both comparisons have practical and theoretical relevance. In this study, we chose texts with abstract topics for which it was easy to find related pictures. Especially for the abstract topics we chose (self-control and population development) it is difficult to find completely unrelated images.

Another possible criticism is that the current study does not provide insight into the phase in the learning process at which the effect occurs. As we explain our pattern of results by mutual facilitation of related content, we assume that the effect has its origin during learning with the text-picture combination. There may be a mutual facilitation of related concepts in working memory that leads to better encoding into long term memory (Davelaar et al., 2006). Beyond the discussed explanation, there are numerous other potential mechanisms how pictures may affect learning that were not focused in the present study, for example processes triggered when pictures are presented both at learning and at test (Lindner et al., 2021; Schneider et al., 2020). Our design and material make it highly unlikely that these processes, that are initiated at retrieval, influence performance in our experiments. Similarly, effects that are triggered if learning behavior is depicted seem unlikely for the current study (e.g., Mikheeva et al., 2021).

Although our study did not show an effect of the used related decorative pictures on interest measures, our results do not imply that related decorative pictures are per se incapable of increasing interest in text or topic. Using different text-picture combinations or more pictures might help to increase interest in the learning content.

## Future research directions

The assumption that beneficial effects of related decorative pictures are caused by mutual maintenance of the activation of related concepts leading to deeper encoding (Davelaar et al., 2006) should be addressed by future research. A *mutual* facilitation based on spreading activation processes, should lead to a positive effect of relatedness not only on learning text content but also on the retention of picture content. Therefore, a logical next step is also testing memory for pictures in addition to memory for text content in a design including the contrast between related and unrelated pictures to gain further insight into the underlying processes. In a within participants design it may further be interesting to correlate the effect of related pictures with working memory performance. As mutual facilitation processes extend the working memory capacity (e.g., Kowialiewski et al., 2021) the performance benefit should be more pronounced for participants with low working memory capacity.

Related decorative pictures did not influence the bias measure. At first glance, this seems to contradict findings of a multimedia heuristic that is inadequately increased confidence after learning with decorative pictures, which may result in an illusion of learning (e.g., Serra & Dunlosky, 2010; Wiley, 2019). A potential explanation for this difference in results is that, in our studies and in contrast to most previous studies observing a multimedia heuristic, we measured metacognitive monitoring at test that is, retrospectively. In this case, metacognitive judgments are based on actual retrieval attempts. In contrast, prospective metacognitive judgments of learning are not informed by the experience of delayed recall attempts. To what extent this difference accounts for detecting the multimedia heuristic could be addressed by further research.

Returning to the starting point, this study clearly demonstrated that and why decorative pictures are not necessarily seductive in nature and that they can promote learning beyond explaining text contents.

**Supplementary Information** The online version of this article contains supplementary material available <https://doi.org/10.1007/s11251-023-09618-8>.

**Funding** Open Access funding enabled and organized by Projekt DEAL.

## Declarations

**Conflict of interest** The authors report there are no competing interests to declare.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

## References

- Anderson, J. R. (1976). *Language, memory, and thought*. Erlbaum.  
Anderson, J. R. (1983). *The architecture of cognition*. Erlbaum.

- Anderson, J. R. (1993). *Rules of the mind*. Erlbaum.
- Barenberg, J., & Dutke, S. (2019). Testing and metacognition: Retrieval practise effects on metacognitive monitoring in learning from text. *Memory*, 27(3), 269–279. <https://doi.org/10.1080/09658211.2018.1506481>
- Barenberg, J., & Dutke, S. (2021). Retrieval practice effects in a psychology lecture: Illustrating the relevance of study design, item difficulty, and selection of dependent measures. *Psychology Learning & Teaching*. <https://doi.org/10.1177/14757257211049312>
- Baron, R. M., & Kenny, D. A. (1986). The moderator–mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. *Journal of Personality and Social Psychology*, 51(6), 1173–1182. <https://doi.org/10.1037/0022-3514.51.6.1173>
- Bartsch, R. A., & Cobern, K. M. (2003). Effectiveness of powerpoint presentations in lectures. *Computers & Education*, 41(1), 77–86. [https://doi.org/10.1016/S0360-1315\(03\)00027-7](https://doi.org/10.1016/S0360-1315(03)00027-7)
- Baumeister, R. F. (01.2018). Wo ein Wille ist ... [Where there is a will ...]. *Spektrum der Wissenschaft: Spektrum Kompakt*, 2018(01 Motivation und Kreativität), pp. 21–27.
- Baumeister, R. F., Bratslavsky, E., Muraven, M., & Tice, D. M. (1998). Ego depletion: Is the active self a limited resource? *Journal of Personality and Social Psychology*, 74(5), 1252–1265. <https://doi.org/10.1037//0022-3514.74.5.1252>
- Baumeister, R. F., Muraven, M., & Tice, D. M. (2000). Ego depletion: A resource model of volition, self-regulation, and controlled processing. *Social Cognition*, 18(2), 130–150. <https://doi.org/10.1521/soco.2000.18.2.130>
- Bethke, J. (2009). *Praxis Geographie – Ausgabe für die S1 in Nordrhein-Westfalen Schülerband 3 [Practical geography - Version for S1 in North Rhine-Westphalia Student Edition 3]*. Westermann.
- Bösch, M. (2012). *Seydlitz - Erdkunde [Seydlitz - Geography]*. Schroedel.
- Bremm, A., & Latz, W. (2012). *Diercke - Erdkunde [Diercke - Geography]* (Nordrhein-Westfalen, Realschule, Sekundarschule). Westermann.
- Butcher, K. R. (2014). The multimedia principle. In R. E. Mayer (Ed.), *The cambridge handbook of multimedia learning* (pp. 174–205). Cambridge University Press. <https://doi.org/10.1017/CBO9781139547369.010>
- Claaßen, K. (2015). Praxis Blatt Japan – Seniorenwindeln schlagen Babywindeln [Practice sheet Japan - senior diapers beat baby diapers]. *Praxis Geographie*, 10, 53–54.
- Claaßen, K. (2016). Bevölkerungspolitik in China. Zum Ende der „Ein-Kind-Politik“ [Population policy in China. On the end of the "one-child policy"]. *Praxis Geographie aktuell*(5). <https://www.westermann.de/anlage/4582318/Bevoelkerungspolitik-in-China-Zum-Ende-der-Ein-Kind-Politik>
- Clark, J. M., & Paivio, A. (1991). Dual coding theory and education. *Educational Psychology Review*, 3(3), 149–210. <https://doi.org/10.1007/BF01320076>
- Davelaar, E. J., Goshen-Gottstein, Y., Ashkenazi, A., Haarmann, H. J., & Usher, M. (2005). The demise of short-term memory revisited: Empirical and computational investigations of recency effects. *Psychological Review*, 112(1), 3–42. <https://doi.org/10.1037/0033-295X.112.1.3>
- Davelaar, E. J., Haarmann, H. J., Goshen-Gottstein, Y., & Usher, M. (2006). Semantic similarity dissociates short- from long-term recency effects: Testing a neurocomputational model of list memory. *Memory & Cognition*, 34(2), 323–334. <https://doi.org/10.3758/BF03193410>
- Dunlosky, J., & Nelson, T. O. (1992). Importance of the kind of cue for judgments of learning (jol) and the delayed-jol effect. *Memory & Cognition*, 20(4), 374–380. <https://doi.org/10.3758/BF03210921>
- Dutke, S., & Barenberg, J. (2015). Easy and informative: Using confidence-weighted true–false items for knowledge tests in psychology courses. *Psychology Learning & Teaching*, 14(3), 250–259. <https://doi.org/10.1177/1475725715605627>
- Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G\*power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, 39(2), 175–191. <https://doi.org/10.3758/bf03193146>
- Frank, F., Brodengeier, E., & Rausch, C. (Eds.). (2009). *Terra - Erdkunde Gymnasium Nordrhein-Westfalen [Terra - Geography Gymnasium North Rhine-Westphalia]*. Klett.
- Harp, S. F., & Mayer, R. E. (1997). The role of interest in learning from scientific text and illustrations: On the distinction between emotional interest and cognitive interest. *Journal of Educational Psychology*, 89(1), 92–102. <https://doi.org/10.1037/0022-0663.89.1.92>
- Hayes, A. F. (2017). *Introduction to mediation, moderation, and conditional process analysis: A regression-based approach* (2nd ed.). Guilford Publications.
- Hidi, S., & Renninger, K. A. (2006). The four-phase model of interest development. *Educational Psychologist*, 41(2), 111–127. [https://doi.org/10.1207/s15326985Sep4102\\_4](https://doi.org/10.1207/s15326985Sep4102_4)

- Hines, J. C., Touron, D. R., & Hertzog, C. (2009). Metacognitive influences on study time allocation in an associative recognition task: An analysis of adult age differences. *Psychology and Aging, 24*(2), 462–475. <https://doi.org/10.1037/a0014417>
- Hoppe, W., Jonas, K., & Sesemann, O. (2018). Tansanias Chancen auf eine demographische Dividende. Das Agenda 2030-Ziel Nr. 4 „Hochwertige Bildung“ [Tanzania's opportunities for a demographic dividend. Agenda 2030 Goal No. 4 “Quality Education”]. *Praxis Geographie, 5*, 34–39.
- Jaeger, A. J., & Wiley, J. (2014). Do illustrations help or harm metacomprehension accuracy? *Learning and Instruction, 34*(6), 58–73. <https://doi.org/10.1016/j.learninstruc.2014.08.002>
- Kowialiewski, B., Lemaire, B., & Portrat, S. (2021). How does semantic knowledge impact working memory maintenance? Computational and behavioral investigations. *Journal of Memory and Language, 117*(1), 104208. <https://doi.org/10.1016/j.jml.2020.104208>
- Lenzner, A., Schnotz, W., & Müller, A. (2013). The role of decorative pictures in learning. *Instructional Science, 41*(5), 811–831. <https://doi.org/10.1007/s11251-012-9256-z>
- Levie, W. H., & Lentz, R. (1982). Effects of text illustrations: A review of research. *ECTJ, 30*(4), 195–232. <https://doi.org/10.1007/BF02765184>
- Leopold, C., Mayer, R. E., & Dutke, S. (2019). The power of imagination and perspective in learning from science text. *Journal of Educational Psychology, 111*(5), 793–808. <https://doi.org/10.1037/edu0000310>
- Lindner, M. A., Eitel, A., Barenthien, J., & Köller, O. (2021). An integrative study on learning and testing with multimedia: Effects on students' performance and metacognition. *Learning and Instruction, 71*(1), 101100. <https://doi.org/10.1016/j.learninstruc.2018.01.002>
- Little, J. L., & McDaniel, M. A. (2015). Metamemory monitoring and control following retrieval practice for text. *Memory & Cognition, 43*(1), 85–98. <https://doi.org/10.3758/s13421-014-0453-7>
- Magner, U. I., Schwonke, R., Aleven, V., Popescu, O., & Renkl, A. (2014). Triggering situational interest by decorative illustrations both fosters and hinders learning in computer-based learning environments. *Learning and Instruction, 29*(1), 141–152. <https://doi.org/10.1016/j.learninstruc.2012.07.002>
- Mayer, R. E. (1989). Systematic thinking fostered by illustrations in scientific text. *Journal of Educational Psychology, 81*(2), 240–246. <https://doi.org/10.1037/0022-0663.81.2.240>
- Mayer, R. E. (2022). Cognitive theory of multimedia learning. In R. E. Mayer (Ed.), *The cambridge handbook of multimedia learning* (pp. 57–72). Cambridge University Press. <https://doi.org/10.1017/9781108894333>
- Metcalfe, J., & Shimamura, A. P. (1994). *Metacognition: Knowing about knowing*. MIT Press. <https://doi.org/10.7551/mitpress/4561.001.0001>
- Mihalca, L., & Mengelkamp, C. (2020). Effects of induced levels of prior knowledge on monitoring accuracy and performance when learning from self-regulated problem solving. *Journal of Educational Psychology, 112*(4), 795–810. <https://doi.org/10.1037/edu0000389>
- Mikheeva, M., Schneider, S., Beege, M., & Rey, G. D. (2021). The influence of affective decorative pictures on learning statistics online. *Human Behavior and Emerging Technologies, 3*(3), 401–412. <https://doi.org/10.1002/hbe.2.250>
- Paivio, A. (1991). Dual coding theory: Retrospect and current status. *Canadian Journal of Psychology/Revue Canadienne de Psychologie, 45*(3), 255–287. <https://doi.org/10.1037/h0084295>
- Rey, G. D. (2012). A review of research and a meta-analysis of the seductive detail effect. *Educational Research Review, 7*(3), 216–237. <https://doi.org/10.1016/j.edurev.2012.05.003>
- Rohr, M., & Wentura, D. (2022). How emotion relates to language and cognition, seen through the lenses of evaluative priming paradigms. *Frontiers in Psychology, 13*(1068), 3750. <https://doi.org/10.3389/fpsyg.2022.911068>
- Schaffner, E., & Schiefele, U. (2008). Familiäre und individuelle Bedingungen des Textlernens [Family and individual conditions of text learning]. *Psychologie in Erziehung Und Unterricht, 55*(4), 238–252.
- Scherer, D., & Wentura, D. (2018). Combining the post-cue task and the perceptual identification task to assess parallel activation and mutual facilitation of related primes and targets. *Experimental Psychology, 65*(2), 84–97. <https://doi.org/10.1027/1618-3169/a000396>
- Schiefele, U. (1990). Thematisches Interesse, Variablen des Leseprozesses und Textverstehen [Topic interest, variables of the reading process, and text comprehension]. *Zeitschrift für Experimentelle und Angewandte Psychologie, 37*(2), 304–332.
- Schmitz, M., & Wentura, D. (2012). Evaluative priming of naming and semantic categorization responses revisited: A mutual facilitation explanation. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 38*(4), 984–1000. <https://doi.org/10.1037/a0026779>
- Schmitz, M., Wentura, D., & Brinkmann, T. A. (2014). Evaluative priming in a semantic flanker task: Erp evidence for a mutual facilitation explanation. *Cognitive, Affective & Behavioral Neuroscience, 14*(1), 426–442. <https://doi.org/10.3758/s13415-013-0206-2>

- Schneider, S., Dyrna, J., Meier, L., Beege, M., & Rey, G. D. (2018). How affective charge and text–picture connectedness moderate the impact of decorative pictures on multimedia learning. *Journal of Educational Psychology, 110*(2), 233–249. <https://doi.org/10.1037/edu0000209>
- Schneider, S., Nebel, S., Beege, M., & Rey, G. D. (2020). The retrieval-enhancing effects of decorative pictures as memory cues in multimedia learning videos and subsequent performance tests. *Journal of Educational Psychology, 112*(6), 1111–1127. <https://doi.org/10.1037/edu0000432>
- Schneider, S., Nebel, S., & Rey, G. D. (2016). Decorative pictures and emotional design in multimedia learning. *Learning and Instruction, 44*(4), 65–73. <https://doi.org/10.1016/j.learninstruc.2016.03.002>
- Schnotz, W. (2022). Integrated model of text and picture comprehension. In R. E. Mayer (Ed.), *The cambridge handbook of multimedia learning* (pp. 82–99). Cambridge University Press. <https://doi.org/10.1017/9781108894333>
- Schnotz, W., & Bannert, M. (2003). Construction and interference in learning from multiple representation. *Learning and Instruction, 13*(2), 141–156. [https://doi.org/10.1016/S0959-4752\(02\)00017-8](https://doi.org/10.1016/S0959-4752(02)00017-8)
- Schweighöfer, B., & Herre, R. (2017). Welternährung. Globale Landwirtschaft und der Skandal des Hungers [World food. Global agriculture and the scandal of hunger]. *Praxis Geographie*(12), 4–9.
- Schwepe, J., Eitel, A., & Rummer, R. (2015). The multimedia effect and its stability over time. *Learning and Instruction, 38*(4), 24–33. <https://doi.org/10.1016/j.learninstruc.2015.03.001>
- Serra, M. J., & Dunlosky, J. (2010). Metacomprehension judgements reflect the belief that diagrams improve learning from text. *Memory, 18*(7), 698–711. <https://doi.org/10.1080/09658211.2010.506441>
- Souchay, C. (2007). Metamemory in Alzheimer’s disease. *Cortex, 43*, 987–1003. [https://doi.org/10.1016/S0010-9452\(08\)70696-8](https://doi.org/10.1016/S0010-9452(08)70696-8)
- Statista. (2018a). *Weltbevölkerung von 1950 bis 2017 (in Milliarden) [World population from 1950 to 2017 (in billions)]*. <https://de.statista.com/statistik/daten/studie/1716/umfrage/entwicklung-der-weltbevoelkerung/>
- Statista. (2018b). *Zuwachs der Weltbevölkerung, angegeben in unterschiedlichen Zeiteinheiten (Stand 2017) [Increase in world population, given in different time units (as of 2017)]*. <https://de.statista.com/statistik/daten/studie/1816/umfrage/zuwachs-der-weltbevoelkerung/>
- Sundararajan, N., & Adesoro, O. (2020). Keep it coherent: A meta-analysis of the seductive details effect. *Educational Psychology Review, 32*(3), 707–734. <https://doi.org/10.1007/s10648-020-09522-4>
- Thiede, K. W., Anderson, M. C. M., & Theriault, D. (2003). Accuracy of metacognitive monitoring affects learning of texts. *Journal of Educational Psychology, 95*(1), 66–73. <https://doi.org/10.1037/0022-0663.95.1.66>
- Wiley, J. (2019). Picture this! Effects of photographs, diagrams, animations, and sketching on learning and beliefs about learning from a geoscience text. *Applied Cognitive Psychology, 33*(1), 9–19. <https://doi.org/10.1002/acp.3495>
- Wiley, J., Griffin, T. D., & Thiede, K. W. (2005). Putting the comprehension in metacomprehension. *Journal of General Psychology, 132*(4), 408–428. <https://doi.org/10.3200/GENP.132.4.408-428>

**Publisher’s Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.