



Can we further improve tablet-based drawing to enhance learning? An empirical test of two types of support

Steffen P. Schmidgall¹ · Katharina Scheiter^{1,2}  · Alexander Eitel^{1,3}

Received: 19 March 2019 / Accepted: 24 April 2020 / Published online: 11 May 2020
© The Author(s) 2020

Abstract

Digital drawing can foster learning, but only if the drawing is of sufficient quality. Hence, the focus of the present study was to investigate whether and how two types of drawing support may foster drawing quality and, in turn, learning outcomes. To this end, participants ($N=156$) were randomly assigned to one of four conditions, in which they either just read text (control), were prompted to make a free-hand representational drawing (unsupported drawing), or they were additionally supported in their drawing efforts because a background (global support) or single elements for the drawing (local support) were already provided. Learning outcomes were assessed by means of recognition, transfer, and a drawing test. Results revealed that students from all three drawing conditions (unsupported, global, and local support) scored better on the transfer and drawing tests than the control condition. Both types of drawing support did neither increase drawing quality nor learning in comparison to unsupported drawing. Reasons for the latter findings are discussed.

Keywords Learner-generated drawing · Learning with multiple representations · Cognitive load · Drawing aids · Assistance dilemma · Technology-based learning and instruction

Introduction

Making visualizations by means of drawing is an integral part of scientific reasoning and insight (Ainsworth et al. 2011). It can be an effective means of expressing thoughts, of making one's own ideas or ideas of others explicit that may otherwise remain implicit. Moreover, students can be asked to engage in drawing as a strategy to learn from text by visualizing how objects and their relations look like, a process referred to as representational drawing (e.g., Schmeck et al. 2014). This can foster deeper understanding of ideas expressed in a written text, especially if main ideas from the text appear in the drawing

✉ Katharina Scheiter
k.scheiter@iwm-tuebingen.de

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

² University of Tübingen, Tübingen, Germany

³ Department of Psychology, Giessen University, Giessen, Germany

and are arranged in a meaningful manner (see e.g., Fiorella and Mayer 2015; Fiorella and Zhang 2018; Leutner and Schmeck 2014; Van Meter and Garner 2005, for reviews).

In other words, in order to optimally profit from drawing, the self-generated visualization should capture the essence of the text. In the context of drawing for learning, however, this is often not the case. Although representational drawing has been shown to be an effective strategy to learn from written text, students sometimes fail to produce coherent visualizations that capture the texts' main ideas, and hence do not show better learning outcomes (e.g., Scheiter et al. 2017). Accordingly, drawing quality (i.e., the degree to which a drawing provides an accurate representation of the text content) has been shown to be predictive of students' learning outcome when drawing to learn from text, a phenomenon called the prognostic drawing effect (e.g., Schmeck et al. 2014).

Against this backdrop, it has been suggested to support students in their drawing efforts so that they will create higher-quality drawings and, in turn, show better learning outcome (Van Meter and Garner 2005). However, a challenge in designing support for drawing pertains to the assistance dilemma (Koedinger and Alevan 2007). On the one hand, students are likely to fail to produce high-quality drawings with too little support, for instance, when creating free-hand drawings from scratch. On the other hand, drawing with too much support may likewise be ineffective because it constraints the drawing process to a degree where desirable generative activities are eventually no longer necessary. For instance, this may be the case if a learner only needs to complete a pre-given visualization that already comprises most of the main ideas; hence, requiring only little, if any elaboration of the text. Against this backdrop, the aim of the present study was to investigate the effectiveness of two types of drawing support that—while they both were intended to enhance drawing quality—differed in the degree to which they constrained the drawing process. Hence, the aim of the present study was not so much to examine the effectiveness of drawing per se, but to test how this way of learning can be further improved.

In the following, we will first refer to the theoretical reasons for why drawing can be helpful, followed by a section on the role of drawing quality, followed by possibilities on how to improve drawing quality that may, in turn, improve understanding.

How does drawing foster comprehension?

In their Cognitive Model of Drawing Construction (CMDC), Van Meter and Firetto (2013) describe the process of creating representational drawings when learning from instructional text as a set of interleaved perceptual and cognitive processes that contribute to the construction of an internal mental model of the text which is then externalized in a drawing. The CMDC is based on contemporary models of learning with multiple representations such as the integrated model of text and picture comprehension (ITPC; Schnotz 2014).

According to the CMDC (Van Meter and Firetto 2013), when reading a text, verbal information enters the sensory register through the eyes and is then organized into graphemic lexical patterns that are further processed to form a propositional representation. This propositional representation contains the information explicitly mentioned in the text. To achieve deeper understanding of the text, additional information, for instance, regarding underlying implicit relations can be inferred with the help of prior knowledge (Van Dijk and Kintsch 1983), thereby allowing construction of a mental model. A mental model is defined as an internal representation of the subject matter that shares similarities to a mental image (e.g., Johnson-Laird 2005) and allows learners to reason about how a system operates by mentally simulating consequences depending on changes to the system. For

instance, a mental model of a human arm allows reasoning about what would happen if one of the muscles was injured and could not be moved properly. Thus, having an accurate mental model in mind is a prerequisite for answering deep comprehension, transfer questions that require students to move beyond the information that was explicitly given in the instruction (Mayer 2002).

However, constructing a mental model from just text can be effortful and subject to mistakes. It requires transferring information from an abstract descriptive code into a depictive, image-like mental representation. Therefore, students may refrain from constructing a mental model when learning from expository text alone. Rather, their mental representation may be tied to a shallow, propositional level (i.e. when learning about biomechanics: “muscles are connected to the bones through the myofibrils”), whereas an adequate image-like mental model would require them to visualize the elements and their relations (e.g., how muscles are specifically connected to the bones through myofibrils).

According to the CMDC (Van Meter and Firetto 2013), the instruction to generate a representational drawing will prompt students to engage in the cognitive activities that are necessary to construct a mental model, which can then be used to derive a perceptual image that is finally externalized in the drawing. Once externalized on the page, the drawing provides feedback about the coherence and completeness of their mental models that can be improved by reviewing the text contents. Hence, a drawing task can stimulate and improve mental model construction (Van Meter and Firetto 2013) if students actively reconsider and evaluate their drawings. Accordingly, while some aspects of a drawing may be derived directly from a propositional representation of the text, a mental model reflecting deeper comprehension is seen both as a prerequisite and consequence of drawing.

In line with this reasoning, empirical research has shown that a drawing task was beneficial for “higher-order” assessments of knowledge such as comprehension (Alesandrini 1981), transfer (Leopold and Leutner 2012; Schmidgall et al. 2019; Schwamborn et al. 2010, 2011), or visuo-spatial tests (Lansing 1984; Leopold and Leutner 2012; Schmeck et al. 2014; Schmidgall et al. 2019; Schwamborn et al. 2011). It was not beneficial for “lower-order” knowledge assessments such as recognition (Van Meter et al. 2006) and factual knowledge (Leutner et al. 2009; Snowman and Cunningham 1975; Van Meter 2001). These results are in line with CMDC, assuming that beneficial effects of drawing appear only if the type of assessment is matched to the type of mental model that is constructed through the drawing activity (Van Meter and Firetto 2013).

Why does drawing not always foster comprehension? The role of drawing quality

According to Van Meter and Garner (2005) for easy learning tasks drawing quality should be a direct consequence of text comprehension, whereas for more challenging tasks drawing can actually improve comprehension. Many studies have highlighted the strong relation between the accuracy of drawings generated during learning from text and the level of comprehension as measured afterwards (prognostic drawing effect, Schmeck et al. 2014; Schwamborn et al. 2010). In particular, the ineffectiveness of drawing instructions to enhance learning found in some studies could be explained by the low-quality drawings that learners had produced (Leopold et al. 2013; Leutner et al. 2009). Likewise, within groups with drawing tasks, learners who constructed high-quality drawings tended to score higher on learning outcome tests than learners who construct low-quality drawings (Greene 1989; Leutner et al. 2009; Mason et al. 2013; Scheiter et al. 2017; Schmeck et al. 2014; Schmidgall et al. 2019; Van Meter 2001; Van Meter et al. 2006).

There may be at least two reasons for why students create low-quality drawings and hence do not show better comprehension when drawing to learn. First, their cognitive capacities may be overtaxed by the difficulty of the learning material. Following cognitive load theory (e.g. Paas et al. 2003; Sweller 2005), when students read text about a complex topic (e.g., about biomechanics) for which they do not have much prior knowledge, they already have a high (intrinsic) cognitive load and not many working memory resources left to engage in further stimulating activities such as drawing. Hence, whereas activities such as drawing are potentially beneficial to comprehension, they may be detrimental when there is the risk of cognitive overload. Therefore, although not intended, a (free-hand) drawing task may additionally burden working memory without stimulating processes that are directly relevant for the learning goal, thereby increasing extraneous cognitive load (i.e., reflected by means of perceived task difficulty; Cierniak et al. 2009; Schmeck et al. 2014) that is undesirable for learning.

Second, students may struggle to manage the logistics of drawing (Leutner et al. 2009); for instance, when they are concerned with drawing aesthetically pleasing pictures in addition to having to visualize the main concepts and their relations from text. These struggles may have their origin in a more substantial challenge, namely, the need to mentally imagine complex visuo-spatial objects and their relations. According to Kosslyn (1994) in order to generate a perceptual image, which can then be externalized, image files need to be retrieved from long-term memory. These image files are the building blocks of the image (e.g., a single muscle fiber). They can be manipulated (i.e., scanned, rotated, transformed, zoomed) within the spatial medium, a mental space with limited capacity, which allows for composing the mental image based on the individual image files. Once composed, the image can be inspected by zooming into it, thereby changing the level of resolution or detail that can be accessed from the mental image. To create a perceptual image that corresponds to the mental model, the relevant image files have to be retrieved and then arranged so that they will accurately reflect the mental model. This process of creating a perceptual image may be particularly difficult, if an image consists of many, interconnected image files and if the spatial orientation of these image elements is paramount to the functioning of the imagined system. In the latter case, many transformations to the original image files may be necessary to compose the overall image. Due to the limited capacity of the spatial medium, it may be difficult for a learner to retrieve the resulting overall image from memory as a whole, thereby yielding an incomplete and inaccurate drawing.

As a consequence, measures to support drawing should be tailored towards reducing the demands associated with creating a drawing, while at the same time still engaging learners in generative learning activities.

How to support drawing to foster comprehension

Students should receive only as much support as they need to provide good drawings (e.g., Van Meter 2001; Van Meter et al. 2006); drawing with too much support may yield limited benefits because students will not engage in generative activities anymore. One is, thus, faced with an assistance dilemma (Koedinger and Alevan 2007), stating that while assisting performance (here: drawing quality) during instruction can sometimes improve learning, sometimes making performance more difficult during instruction improves learning (e.g., Bjork and Bjork 2011). A certain difficulty, such as the need to generate one's own drawing, can be desirable for learning.

The way drawing has been implemented in research varies from unsupported free-hand drawing (e.g., Leopold and Leutner 2012; Scheiter et al. 2017; Schmidgall et al. 2019) to highly supported drawing, where the drawing could be constructed by assembling pre-defined visual elements (e.g., Schwamborn et al. 2010). Most of the studies on drawing have contrasted these highly variable drawing conditions with a read-only control group or conditions, where students had to perform other text-based learning strategies such as writing a summary (Leopold and Leutner 2012) or self-explaining (Scheiter et al. 2017) in order to determine whether drawing is effective. Only very few studies have compared different types of drawing conditions that varied in the level of support provided for drawing (see Van Meter 2001, for an exception), which is the focus of the present paper. That is, we are not interested so much in the effectiveness of drawing per se; rather, we are interested in whether by providing support drawing as a learning strategy can be further enhanced (Fiorella and Zhang 2018). In particular, our focus is on supporting processes that are related to deriving a perceptual image and externalizing it.

Referring back to Kosslyn's mental imagery theory (1994), two types of support can be distinguished in this regard: First, support can be tailored towards helping students with the identification and externalization of single elements within a drawing by providing them with these elements (e.g., by offering a list of pre-drawn shapes). This type of local support relieves students from the need to retrieve the correct image files from long-term memory when imagining and externalizing their mental model, thereby making it unnecessary to reason about the visual appearance of relevant objects. Second, support can be located at the global level, when it aims at helping students arrange the single elements into a coherent image (i.e., image composition). Even though also this type of support often consists of providing an image file (e.g., a drawing of a scenery or background to which single elements can be added), its function is to help students locate the single elements within the overall image that is supposed to be put together within the spatial medium (cf. Kosslyn 1994). Moreover, providing a global frame relieves students from the need to make decisions about the size of the image and its elements as well as its resolution, thereby further facilitating the translation of the content of the internal spatial medium onto paper or a digital display.

Schmeck et al. (2014) and Schwamborn et al. (2010) investigated the effects of drawing to learn for students at the secondary level of the highest track in Germany in a science domain. Their drawing strategy comprised support by means of a two-part drawing scaffold. First, it included a partially pre-drawn background for students' drawing which students were asked to complete (global support). Importantly, the most relevant ideas from text were not yet provided but were to be completed by the learner. Second, the scaffold included a toolbar showing all the relevant pictorial elements and labels for the drawing, from which learners could choose in order to construct their drawings (local support). No irrelevant elements were provided. Schwamborn et al. (2010) found this way of supported drawing to be effective for science learning when compared to a read-only control group. Schmeck et al. (2014) were able to replicate these findings; moreover, they showed that supported drawing was more effective than learning from instructor-generated illustrations or combinations of drawings and instructor-generated illustrations.

There are at least two open questions regarding the benefits of drawing that is supported at either the global or local level. First, it is yet unclear whether the way drawing was implemented in the aforementioned studies would also be superior to free-hand drawing, which is paramount to considering it a support measure. That is, the aforementioned studies say something only about the effectiveness of (supported) drawing per se. Moreover, it is yet an open question as to which type of support, on a global or a local level, mainly

drove the beneficial effects. According to Kosslyn (1994), both types of processes that may be addressed through these support measures (i.e., retrieval of image files vs. composition of image files within the spatial medium) equally contribute to mental imagery. Hence, it is not possible to derive from this theory a hypothesis as to which support measure will be more effective in the context of drawing.

Present research and hypotheses

In the present study,¹ university students were asked to learn about a science topic (i.e., bio-mechanics of swimming behavior) as this is a typical visuo-spatial learning tasks for which students may benefit from a drawing task (cf. Ainsworth et al. 2011). Drawing was implemented on a tablet computer. Within this context, we investigated whether and how global or local support may affect representational drawing and learning outcomes by assigning participants to one of four conditions. In the first condition (control), students were prompted to read text, meaning to identify the key concepts and to reread passages they did not understand. In the second condition, besides reading students were asked to create free-hand drawings for the text passages without support, meaning they had to decide which concepts should be visualized, how they should look like and be spatially arranged without any frame of reference. In the third condition, the background was already settled for students' drawing (global support), meaning they had to interpret the pre-specified picture (background), decide which concepts should further be visualized, how they should look like and be spatially arranged within the frame already provided. In the fourth condition, in the local support condition, students read text and were asked to create a drawing by selecting relevant elements and labels from a menu bar, which also included distractors, and arranging them in space. We decided to include distractors in the menu bar to keep students from a trial-and-error approach that would simply involve putting all elements somewhere without engaging in deeper thinking. Notably, our sample comprised university students rather than students at the secondary level as in the studies by Schwamborn et al. (2010) and Schmeck et al. (2014); hence, we were confident that our students would be able to cope with the additional demands of selecting relevant elements and would benefit from the need to reason about their adequacy.

In the following, we will derive two main hypotheses (with two sub-hypotheses each) regarding expected differences between these four conditions. As drawing is expected to foster higher-order knowledge in particular (i.e., comprehension, transfer, diagram completion), the hypotheses will refer only to this type of knowledge assessment (cf. Leutner and Schmeck 2014). For lower-order knowledge (recognition), no substantial differences between conditions are expected, which is why no explicit hypotheses were derived.

- According to Hypothesis 1a, all three drawing conditions should outperform the no-drawing control condition, because drawing should stimulate mental model construction that is, in turn, related to better comprehension.
- According to Hypothesis 1b, better drawing quality should be positively associated with learning outcomes across the three conditions (cf. prognostic drawing effect;

¹ The study reported in this paper is Study 3 in the dissertation of Steffen Schmidgall, which is available here for further reference: <https://doi.org/10.15496/publikation-18678>.

Leutner and Schmeck 2014; Van Meter and Firetto 2013). Conversely, drawing should foster learning especially when drawings are of high quality.

- Following Hypothesis 2a, the conditions with global and local drawing support should outperform the condition with unsupported free-hand drawing. The question of which of the two support conditions works best is exploratory.
- Following Hypothesis 2b, global and local drawing support should foster learning via increased drawing quality and lower perceived difficulty. The more the drawing activity is focused on central issues from the text, which is defined as drawing quality, the better they are understood. Moreover, the fewer difficulties students perceive when managing the drawing process (and hence the lower the extraneous cognitive load), the lower might be the risk of a cognitive overload that would be detrimental to performance.

Method

Participants and design

Participants were 157 undergraduate students from a university in the southern part of Germany, who took part in the study voluntarily for an incentive of 12 euros or course credit. Ethics approval for the current study was obtained from the local ethics committee of the Leibniz-Institut für Wissensmedien Tübingen (number: 2014/053). Informed consent was provided by all participants. The number of participants was determined a priori by conducting a power analysis using G*Power 3.1 (Erdfelder et al. 2009) with the medium effect size partial $\eta^2=0.09$, a power of 0.90, and alpha=0.05 resulted in a recommended sample size of 148 participants. Biology, medicine, physics and sports sciences majors were excluded from the experiment because we expected them to have too much prior knowledge on the learning content. One participant, whose German language proficiency seemed not to be sufficient to fully understand the instructions and materials, was excluded from the analyses. Hence, we used the reduced sample of 156 participants (119 female; $M=21.58$ years, $SD=2.94$) to conduct the analyses. Participants were randomly assigned to one of four experimental conditions that resulted from a one-factorial between-subjects design. The four experimental conditions were the non-drawing control group ($n=39$), no support ($n=39$), global support ($n=39$), and local support ($n=39$). Three participants terminated the study prematurely after the transfer test and did not work on the drawing test. These participants were initially assigned to the global support ($n=2$) and local support conditions ($n=1$). So, a reduced sample of participants ($N=153$) was used for the analysis of drawing test performance.

Materials

The learning materials and tests were used in a previous study conducted by Schmidgall et al. (2019), where they proved suitable to demonstrate differences between different visualization-related learning activities (e.g., learning from author-generated illustrations, mental imagery, drawing). For the learning episode, we developed a text about biomechanics in human swimming behavior that contained 983 words in nine paragraphs. Each paragraph was presented on a separate page. The paragraphs were about muscle contraction (112 words), series connection of muscle fibers (98 words), parallel connection of muscle fibers (93 words), pennate muscles (116 words), buoyancy (124 words), the human body

in water (112 words), fluid resistance (113 words), the drive concept action-reaction (87 words) and the drive concept hydrodynamic lift (128 words). The expository text was created by the first author of the present paper through referring to multiple textbooks and WWW resources and cross-validating its key ideas.

The layout of the pages depended on experimental condition (see Fig. 1). In the control condition, only text was presented in the middle of each page aligned to the top. In the other three drawing conditions, the text was presented on the left of each page aligned to the top. In the no support (free-hand drawing) condition (Panel A), the right side of each page was blank with a prompt asking participants to construct their drawing there. In the global support condition (Panel B), the right side of each page contained pre-given partial drawings and a prompt asking participants to complete these drawings. In the no support and global support conditions, participants engaged in actual drawing of lines to create representational drawings of what was described in the written text. In the local support condition (Panel C), each page of the learning material was divided into three parts. The text was presented on the left side and two boxes were presented on the right side of each page, one below the other. The upper box contained different objects to be used to construct drawings. The lower box contained different labels to be used to label the drawings. These boxes contained target items to be used to construct the respective drawings but also distractor items that were useless for constructing correct drawings. Via drag and drop, participants were able to select objects and labels from the boxes and pull them in the middle of the screen to construct drawings on a white canvas. Participants were able to use all objects and all labels that were presented in the boxes on the right side of each page. Objects and labels were replaced after selection and could be used arbitrarily often. Participants received no feedback on whether they selected the correct objects or labels. In the local support condition, participants did not engage in actual drawing of lines to create drawings. Instead, they engaged in selecting processes when choosing objects from the boxes. Then, they arranged the selected objects to a visuo-spatial representation of what was described in the written text. We designed the local support condition to serve two purposes: (1) we included pre-given elements to serve as a constraint function and to reduce the need to reason about the visual appearance of any part of the drawings, and (2) we added distractors to foster selecting processes assuming that reasoning about which elements are correct and which are incorrect should increase deeper elaboration.

In total, nine drawings had to be generated by students in each of the three drawing conditions, that is, one for each text paragraph.

Measures

The same measures were used by Schmidgall et al. (2019) in a previous study.

Control variables

We assessed several control variables to assure comparability between experimental conditions prior to participating in the study. Beyond registering participants' demographics (gender and age), we assessed four concepts that seemed relevant to the present task.

First, we assessed participants' subjective ability to draw and their enjoyment regarding drawing on a 7-point Likert-scale ranging from *not at all* to *very much*.

Second, prior knowledge was measured using 20 verification items based on the text used in the learning episode (e.g. "Muscles are connected to the bones through the

*Panel A***Drive concepts: Action-Reaction**

If the palm of the hand is dipped into the water at an angle of 90° to the water surface, the palm of the hand experiences a resistance in the water, as with the arm pull of the crawl swim. The float can "repel" itself from this resistance. This type of drive follows Newton's third axiom: Action = Reaction:

If the palm of the hand exerts a force on the water (action), a force of the same magnitude but opposite to that of the water acts on the palm of the hand (reaction). The swimmer can therefore "repel" himself from the water with the help of this force, which gives the swimmer propulsion.

Make a drawing that will help you better understand what you have learned.

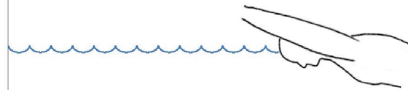
Please check whether your drawing contains all relevant elements.

*Panel B***Drive concepts: Action-Reaction**

If the palm of the hand is dipped into the water at an angle of 90° to the water surface, the palm of the hand experiences a resistance in the water, as with the arm pull of the crawl swim. The float can "repel" itself from this resistance. This type of drive follows Newton's third axiom: Action = Reaction:

If the palm of the hand exerts a force on the water (action), a force of the same magnitude but opposite to that of the water acts on the palm of the hand (reaction). The swimmer can therefore "repel" himself from the water with the help of this force, which gives the swimmer propulsion.

Complete the drawing so that it contains all relevant elements from the text.



Please check whether your drawing contains all relevant elements.

Fig. 1 Example of a page of the learning material (translated from the German original) in the drawing conditions depending on the experimental condition: **a** no support, **b** global support, **c** local support

Panel C

Drive concepts: Action-Reaction

If the palm of the hand is dipped into the water at an angle of 90° to the water surface, the palm of the hand experiences a resistance in the water, as with the arm pull of the crawl swim. The float can "repel" itself from this resistance. This type of drive follows Newton's third axiom: Action = Reaction.

If the palm of the hand exerts a force on the water (action), a force of the same magnitude but opposite to that of the water acts on the palm of the hand (reaction). The swimmer can therefore "repel" himself from the water with the help of this force, which gives the swimmer propulsion.

The diagram shows a hand dipping into water, with arrows indicating the action force (downward) and the reaction force (upward). A legend lists the following terms: Auftrieb, Actio, Eintauchwinkel, Druck, Abstoßen, Schwimmer, Wasser, Armbewegung, Widerstand, and Reactio.

Fig. 1 (continued)

myofibrils.”; Cronbach’s $\alpha=0.67$). Each item consisted of a short statement that was a non-verbatim version of the learning text with a three-choice answer format (‘right’ vs. ‘wrong’ vs. ‘I don’t know the answer’). Participants were rewarded with one point for each correct answer and the resulting scores were transformed into a percentage.

Third, verbal ability was assessed using a computer-based adaption of the sentence finishing subtest of the I-S-T 2000 R (Liepmann et al. 2007) containing 20 multiple-choice items (Cronbach’s $\alpha=0.69$). In this test, participants have to choose one out of four words that fits the meaning of a cloze sentence. One point is awarded for each correct response (cf. Liepmann et al. 2007).

Fourth, spatial ability was assessed with 10 multiple-choice items (Cronbach’s $\alpha=0.65$) taken from the Paper Folding Test (PFT; Ekstrom et al. 1976). In this test, participants have to imagine the folding and unfolding of pieces of paper with holes punched into the folded paper. One item shows the folding of a piece of paper and then holes are punched into it. Participants have to choose the unfolded paper that matches the folded test stimulus out of five alternatives. They receive one point for correct responses, one point is subtracted for an error (cf. Ekstrom et al. 1976).

Dependent variables

Learning outcome, perceived difficulty of the learning material, and drawing quality of constructed drawings during learning were assessed as dependent variables.

First, learning outcome was assessed with a recognition test, a transfer test, and a drawing test. The recognition test consisted of the same 20 verification items that were used for assessing prior knowledge and measured participants’ recognition of the learning content (Cronbach’s $\alpha=0.72$). They referred to the key ideas explained in the text as established by the authors. The transfer test was designed to measure participants’ transfer knowledge and consisted of nine multiple-choice items (Cronbach’s $\alpha=0.63$). An example (translated from the German original) is: ‘How can you explain the lift of an

aircraft? (a) Due to the inclined position and curved shape of the wings, air molecules are dammed under the wings. Air molecules stream faster over the wings where a vacuum emerges. The wings experience lift due to the suction effect. (b) Due to the inclined position and curved shape of the wings, air molecules are dammed above the wings. Air molecules stream faster under the wings where a vacuum emerges. The wings experience lift due to the suction effect. (c) Due to its inclined position, air molecules are dammed under the aircraft. The air resistance increases and pushes against the bottom of the aircraft, thereby, experiencing lift and propulsion at the same time. (d) Due to its inclined position, air molecules are dammed above the aircraft. Air molecules stream faster under the aircraft where a vacuum emerges. The aircraft experiences lift due to the suction effect'. Participants had to choose one of four alternatives and were awarded with one point for a correct answer and the resulting scores were transformed into a percentage. To solve the transfer test, participants had to apply their knowledge acquired through the learning material to new contexts. For example, participants read a paragraph about buoyancy in water. The corresponding transfer question (see the example item above) required participants to apply that knowledge to another context; in this case it was buoyancy in air.

The drawing test consisted of four drawing items and was designed to assess participants' visuo-spatial knowledge of the learning content (Cronbach's $\alpha = 0.69$). Participants had to answer open-ended questions with a drawing or complete pre-given incorrect or partial drawings, and label parts of pre-given drawings (e.g., "Please draw in the missing fibers of a muscle with series connection."). The possible number of points differed between the items according to the number of main concepts required to respond accurately to the corresponding item. For the drawing test, participants could earn a maximum of 13 points; again, scores were transformed into a percentage. Two raters scored 22% of the data for the drawing test. Inter-rater agreement was high (all Krippendorff's alphas > 0.92) so that one of the raters scored the remaining data.

Second, to assess the quality of the drawings constructed for each paragraph during learning, we first identified main idea units within each paragraph and assessed whether these were correctly incorporated into the drawings. The drawing quality score was computed by assigning one point to participants for each correctly drawn main concept unit, yielding a maximum score of 76 points. Then, we calculated the percentage correct. Two raters scored 21% of the data for drawing quality. Inter-rater agreement was acceptable (all Krippendorff's alphas > 0.84) so that one of the raters scored the remaining data. Drawing quality was assessed only for participants in the three drawing conditions because the control condition did not construct drawings during learning.

Third, we measured extraneous cognitive load in terms of perceived difficulty (cf., Cierniak et al. 2009; Schmeck et al. 2014) using three items (e.g., "How easy or difficult was it for you to learn something about biomechanics in human swimming behavior?"; Paas 1992; Cronbach's $\alpha = 0.81$) that were rated on a 7-point Likert-scale ranging from *very easy* to *very difficult*.

Apparatus

Each participant was provided with a tablet computer (“Apple iPad 4 with an integrated 9.7” multi-touch display) on which all materials and measurements were presented.² The presentation of the measures, except for the drawing test, was web-based. The presentation of the learning material in the control, no support, and global support conditions and the presentation of the drawing test was implemented using the application ‘Notability’ (Ginger Labs Inc., version 5.33), which allows for the construction of drawings. This application contains a toolbar at the top of the screen comprising several functions that participants were able to use while working on the learning material, including ‘undo’ (undo the last step), ‘crayon’ (drawing with a pencil; color and thickness of the drawn lines could be adjusted), ‘highlighting’ (drawing transparent, thicker lines; color and thickness of the lines could be adjusted), ‘eraser’ (erasing what was drawn before with a touch on the relevant position) ‘cut and paste’ (selecting parts of drawn objects to move them across the canvas or to duplicate them with a long touch on the selected object) and ‘scrolling’ (if activated only scrolling through the pages is possible; with the touch of one finger). The pages of the learning material were presented below the toolbar. Each iPad was accompanied by a stylus pen to be used during the experiment. For the presentation of the learning material in the local support condition, we used a Windows-based tablet-computer (“Microsoft Surface Pro 2 with an integrated 10.6” touchscreen) because the application we used to implement the learning material in the local support condition did not run on an iOS-based tablet. The learning environment of the local support condition was programmed using ‘Adobe Flash Builder’ (Adobe Systems, version 4.7). Participants were able to undo their last steps or to delete individual objects on the canvas and to turn pages back and forth.

Procedure

Participants were randomly assigned to one of the four experimental conditions. Upon arrival at the lab, they were seated in a semiprivate cubicle in front of a tablet. Participants were tested in groups with a maximum number of five participants per session. At the beginning, some demographic questions were presented, before the manipulation was carried out: participants received different information about how to work on the learning material, depending on the condition they were assigned to. All participants were told that the learning activity (drawing or just reading) they should engage in has proven to be effective for learning. Additionally, participants in the drawing conditions (no support, global support and local support conditions) were instructed to thoroughly read the learning text, to identify all key concepts, and to construct a drawing for each paragraph that contained all important information of the text and that those drawings did not need to look pretty. Participants in the control condition were instructed to thoroughly read the text, identify all key concepts and to reread passages they did not understand properly. Before participants worked on the learning material, they were provided with a short exercise that was designed to familiarize participants with the respective applications in which the learning material was presented depending on the experimental condition.¹ In the control, no

² Participants in the local support condition were handed a Surface-tablet to work on the exercise and the learning material. After completing the learning episode, these participants switched back to an iPad for the remainder of the experiment.

Table 1 Means and standard deviations for the learning outcome measures and perceived difficulty as a function of the four experimental groups

Group	Type of test									
	<i>n</i>	Recognition (% correct)		Transfer test (% correct)		Drawing test (% correct)			Perceived difficulty (0–6)	
		<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>n</i>	<i>M</i>	<i>SD</i>
Control	39	74.36	12.99	41.03	17.69	37.96	20.10	39	2.37	1.28
No support	39	68.08	12.33	51.85*	18.96	47.54*	22.83	39	2.71	1.20
Global support	39	71.28	13.11	46.15	19.17	49.17*	19.16	37	3.26***	1.08
Local support	39	65.13*	14.35	51.28*	20.96	49.52*	17.22	38	3.03*	1.06

*Statistically significant difference from the control group at $p < 0.05$

***Statistically significant difference from the control group at $p < 0.001$

support, and global support conditions, participants were instructed to try out the different functions of Notability by means of drawing rectangles in different colors, erasing the rectangles, cut out and paste the rectangles, highlighting the instructions of the task and undoing previous steps. In the local support condition, participants were presented with several elements that could be combined to a house and were instructed to draw a house by means of drag and drop. Furthermore, they were instructed to delete elements on the canvas and to undo previous steps in order to try out the respective functions. During the learning episode, participants in the control condition were only able to scroll through the learning material, whereas participants in the no support and global support conditions were able to use all previously mentioned functions of Notability. After completing the self-paced learning episode, participants worked on a questionnaire assessing perceived difficulty and interest, the verbal ability and the spatial ability tests as well as the posttest.

Results

Control variables

Before testing the hypotheses, we examined whether the four experimental groups differed in gender, age, subjective ability to draw, enjoyment regarding drawing, prior knowledge, verbal ability, and spatial ability. A chi-square analysis for gender revealed no significant differences between the groups ($p = 0.88$). For the remaining control variables, we computed a multivariate analysis of variance (MANOVA) with experimental condition as between-subjects factor. It revealed no significant effect of experimental condition, Pillai's trace = 0.145, $F(18,447) = 1.26$, $p = 0.21$, partial $\eta^2 = 0.05$. Thus, the experimental conditions were similar with respect to participants' entry characteristics.

Learning outcomes

Means and standard deviations are reported in Table 1. A summary of results is provided in Table 2.

Table 2 Summary of results

Dependent measure	Significant overall group effect	Significant Bonferroni-adjusted comparisons
Recognition test	Yes	Control > local support
Transfer test	Yes	No support > control Local support > control
Drawing test	Yes	No support > control Local support > control Global support > control
Drawing quality	Yes	No support > local support Global support > local support
Perceived difficulty	Yes	Global support > control Local support > control Global support > no support

Assessments of lower-order knowledge

To test the assumption that a drawing activity does not increase assessments of lower-order knowledge, we conducted a univariate analysis of covariance (ANCOVA) with experimental condition as between-subjects factor, recognition performance as dependent variable, and prior knowledge as covariate. It revealed a statistically significant difference between experimental conditions, $F(3,151)=3.56$, $MSE=174.79$, $p<0.05$, partial $\eta^2=0.07$. Bonferroni-adjusted post-hoc multiple comparisons showed that the control condition scored statistically significantly higher than the local support condition ($SE=2.99$, $p=0.01$). All other differences were not significant (all $p_s>0.23$).

We found no benefits of a drawing activity for recognition. This finding is in line with previous studies showing that a drawing activity does not foster lower-order knowledge but rather supports deeper processing of the learning material resulting in increased comprehension (Leutner et al. 2009; Snowman and Cunningham 1975; van Meter 2001; van Meter et al. 2006). This finding provides further evidence for the assumption that benefits of drawing appear only if the type of assessment matches the type of mental model that is constructed through drawing (Van Meter and Firetto 2013).

Assessments of higher-order knowledge

To test the predictions regarding assessments of higher-order knowledge that all three drawing conditions should outperform the control condition (Hypothesis 1a) and that the conditions with global and local support should outperform the condition with unsupported drawing (Hypothesis 2a), we conducted an ANCOVA with experimental condition as between-subjects factor, transfer performance as dependent variable, and prior knowledge as covariate. Experimental conditions differed statistically significantly, $F(3,151)=2.69$, $MSE=371.58$, $p<0.05$, partial $\eta^2=0.05$. Post-hoc multiple comparisons showed that both the no support ($SE=4.37$, $p=0.02$) and the local support conditions ($SE=4.37$, $p=0.02$) outperformed the control condition. The no support and local support conditions did not differ significantly ($SE=4.37$, $p=0.90$). There were no differences between the global support and the other three conditions (all $p_s>0.19$).

For drawing test performance, an ANCOVA with the covariate prior knowledge revealed that the experimental conditions differed statistically significantly, $F(3,148)=2.88$, $MSE=398.49$, $p<0.05$, partial $\eta^2=0.06$. The no support ($SE=4.52$, $p=0.03$), global support ($SE=4.58$, $p=0.02$), and local support conditions ($SE=4.55$, $p=0.01$) scored better than the control condition, thereby performing equally well (all $p_s>0.68$).

These findings support Hypothesis 1a indicating that a drawing activity stimulates mental model construction which is related to better comprehension of the learning content compared to reading only. Drawing support did not further promote comprehension which is not in line with the assumptions of Hypothesis 2a. Drawing implemented with either global and local support seemed to be as effective as unsupported drawing.

Is there a prognostic drawing effect?

We analyzed the quality of the learner-generated drawings constructed in the three drawing conditions to test Hypotheses 1b and 2b. An exploratory analysis showed that drawing quality scores ranged from 3.33 to 71.72% ($M=28.66\%$, $SD=14.15$). A higher drawing quality score indicates more accurate drawings. In line with the prognostic drawing effect (Schwamborn et al. 2010) and Hypothesis 1b, drawing quality was positively associated with better performance in the recognition ($r=0.42$, $p<0.001$), transfer ($r=0.21$, $p<0.05$), and drawing test ($r=0.28$, $p<0.01$).

This finding provides further empirical evidence for the positive relation between high-quality drawings and learning outcomes (cf., Scheiter et al. 2017; Schmeck et al. 2014; Schwamborn et al. 2010; Van Meter 2001; Van Meter et al. 2006). It emphasizes that it is not just important what learners draw but rather that learners focus on the most important aspects of the learning content and draw them correctly.

Effects of drawing support on drawing quality and perceived difficulty

Does drawing support increase drawing quality?

To test whether drawing support increases drawing quality (Hypothesis 2b), we computed an ANOVA with experimental condition as between-subjects factor and drawing quality as dependent variable. It revealed that drawing quality differed significantly between conditions, $F(2,114)=31.26$, $MSE=131.49$, $p<0.001$, partial $\eta^2=0.35$. Bonferroni adjusted multiple comparisons revealed that the no support condition ($M=36.00\%$, $SD=11.80$) constructed statistically significantly more accurate drawings than the local support condition ($M=16.93\%$, $SD=8.00$, $SE=2.60$ $p<0.001$). The global support condition ($M=33.05\%$, $SD=13.83$) also constructed statistically significantly more accurate drawings than the local support condition ($SE=2.60$, $p<0.001$). The no support and the global support conditions did not differ ($p=0.77$).

This finding does not support Hypothesis 2b. Drawing support did not increase drawing quality and local support even decreased drawing quality. It might indicate that global or local support alone provides not sufficient constraint to help learners construct more accurate drawings. A two-part drawing scaffold similar to the ones used by Schmeck et al. (2014) or Schwamborn et al. (2010) seems to be rather suitable for helping learners to construct more accurate drawings.

Does drawing support lower perceived difficulty?

To test the assumption that drawing support should decrease perceived difficulty (Hypothesis 2b), an ANCOVA with prior knowledge as covariate revealed that the experimental conditions differed statistically significantly with regard to perceived difficulty ratings, $F(3,151)=5.29$, $MSE=1.61$, $p<0.01$, partial $\eta^2=0.10$. Bonferroni-adjusted post-hoc multiple-comparisons showed that the control condition reported statistically significantly lower values of perceived difficulty than the global support ($SE=0.29$, $p<0.001$) and the local support conditions ($SE=0.29$, $p=0.02$). Moreover, the no support condition reported statistically significantly less perceived difficulty than the global support condition ($SE=0.29$, $p=0.01$). All other differences were not significant (all $p_s>0.17$).

Perceived difficulty was negatively correlated with recognition performance ($r=-0.28$, $p<0.001$), but was not significantly correlated with transfer and drawing test performance (both $|r_s|<0.11$, $p_s>0.21$).

The results regarding perceived difficulty were not in favor of Hypothesis 2b assuming that drawing support should decrease perceived difficulty. Instead, global and local support increased perceived difficulty hinting at cognitive overload. The findings indicate that neither global nor local support did reduce extraneous cognitive load. Thus, participants' cognitive resources were overtaxed which is potentially why they did not construct more accurate drawings and did not score higher on the comprehension measures compared to participants who engaged in unsupported drawing. It seems that global support does not provide sufficient guidance to effectively constrain the drawing process. The present findings rather indicate that providing learners with backgrounds for their drawings leads to cognitive overload because learners are concerned with managing to match the drawings emerging from their own mental images to the author-provided background images. Local support as applied by providing elements and distractors for constructing drawings seems to be detrimental with respect to cognitive load. Instead of promoting beneficial selection processes as postulated by theories of multimedia learning (e.g., Mayer 2002, 2014; Schnotz 2014), distractors rather appear to be accompanied by higher cognitive demands for identifying the correct elements for drawings also in adults. This finding is comparable to a previous study conducted with primary school children (Lesgold et al. 1975). First-graders were given cutout figures and were instructed to construct drawings. The drawing strategy did promote learning only when no distractor elements were provided alongside the correct elements. Taken together, the present findings suggest that on the one hand global support alone may not provide sufficient constraints and on the other hand local support with distractors might be too much for learners to benefit more from drawing.

Moreover, perceived difficulty ratings seem to suggest that unsupported drawing promotes mental offloading through external processing which is in line with the assumption that external representations are assumed to reduce the amount of cognitive effort which is required to solve a certain task (Scaife and Rogers 1996). This is consistent with previous studies showing that a drawing activity is associated with a similar amount of extraneous cognitive load than a text-focused learning activity (Schmeck et al. 2014; Schwamborn et al. 2011).

Discussion

We investigated whether it is more beneficial to draw with support compared to drawing without support and which type of drawing support, on a global or on a local level, is more helpful. To this end, we compared a non-drawing control condition with an unsupported free-hand drawing condition and two drawing conditions comprising support. The global support condition received pre-drawn backgrounds in order to prompt drawing of the relevant elements without having to engage in unnecessary reasoning regarding the overall arrangement of visual elements on the screen (i.e. on where to start on the page or the size of the overall image and its elements). The local support condition received different elements and labels to choose from in order to construct drawings. This was expected to free learners from reasoning about the visual appearance of the single elements, thereby enabling them to focus on the more meaningful spatial relations between these elements. The two types of support had been implemented together as components of an effective drawing to learn strategy in previous studies (Schmeck et al. 2014; Schwamborn et al. 2010), where they had been proven beneficial compared with a read-only control group. The present paper aimed at disentangling the effects of these two approaches and investigating whether they would serve as a tool to support drawing also when compared with a free-hand drawing condition.

In brief, we replicated previous findings in that drawing promoted the acquisition of higher-order knowledge. Moreover, a higher quality of constructed drawings was associated with better test performance. However, contrary to our predictions none of the implemented types of support further promoted learning when compared with an unsupported drawing condition. Rather, the three drawing conditions were equally effective in promoting learning compared with a read-only control group no matter whether and how additional support had been implemented. Moreover, support did neither increase drawing quality nor reduce perceived difficulty. Taken together, the results show that the two support measures were not suited to further enhance drawing to learn.

How much support for drawing is enough but not too much?

Previous research on drawing indicates that some kind of support during drawing is necessary in order for learners to benefit more from a drawing activity (Alesandrini 1981; Lesgold et al. 1975, 1977; Schmeck et al. 2014; Schwamborn et al. 2010; van Meter 2001; van Meter et al. 2006). We were interested whether different types of support are equally beneficial. Unexpectedly, our results showed that none of the applied support, neither on a global nor on a local level, promoted learning beyond a benefit of unsupported drawing, thereby contradicting previous research on instructional support during drawing (for overviews see Leutner and Schmeck 2014; van Meter and Firetto 2013; van Meter and Garner 2005). The findings regarding drawing quality and perceived difficulty indicate that global support was not sufficiently constrained for learners to facilitate the mechanics of constructing drawing, whereas local support containing distractors was too much constrained, both resulting in cognitive overload. Nevertheless, university students were still able to benefit more from drawing with support compared with the read-only control group, which is in line with results from Schwamborn et al. (2010) and Schmeck et al. (2014).

There are at least two open questions that require future research: First, it could be that global and local support enhance drawing to learn only when presented together as one

coherent drawing strategy as was the case in the aforementioned studies. From a theoretical point of view, they address different processes during creating and externalizing perceptual images (cf. Kosslyn 1994), but this does not exclude the possibility that they might work only when both of these processes are supported simultaneously. In a study by Alesandrini (1981) verbal prompts were used to guide students to paying attention to either local (visual appearance of elements) or global (overall visuo-spatial arrangement and composition) aspects of their drawing. The best learning gains were obtained when students were guided to paying attention to more global, holistic aspects of their drawings. This provides tentative evidence that there may be conditions under which a global support strategy is more effective for learning.

Second, further research also needs to investigate whether university students might require less support than the secondary-level students investigated in the aforementioned studies by Schwamborn et al. (2010) and Schmeck et al. (2014). Our results do not suggest ceiling effects for either learning outcomes or drawing quality; however, it could be that to improve university students' performance other instructional support measures are needed than the ones investigated in the present study, which pertained to supporting the creation and externalization of a perceptual image.

The relation between drawing quality and learning outcomes

We were interested in replicating the prognostic drawing effect (Schwamborn et al. 2010) which describes the finding that a higher quality of constructed drawings is associated with better test performance (e.g., Scheiter et al. 2017; Schmeck et al. 2014; Schwamborn et al. 2010; van Meter 2001; van Meter et al. 2006). Our results replicate previous findings regarding the prognostic drawing effect, revealing a positive relation between drawing quality and learning outcome performance.

Furthermore, we expected both types of drawing support to positively influence drawing quality because support should assist learners in constructing more accurate drawings. Contrary to our predictions, support resulted in drawings of similar (global support) or even worse quality (local support) than drawing without support. The results indicate a trade-off between providing support and learners' cognitive engagement during learning. While higher cognitive engagement is assumed to promote learning (Chi 2009; Chi and Wylie 2014), it seems that drawing aids reduced cognitive engagement reflected by lower drawing quality, higher ratings of perceived difficulty, and equal learning compared with unsupported drawing.

Limitations

The text that was used in the learning episode was relatively short. It might be useful to conduct future studies with longer and more ecologically valid texts (such as content that university students must study as part of a course or that pupils must study as part of their curriculum). Moreover, a replication of the present study with pupils or even in the classroom may be advisable to elaborate the present findings.

It could be argued that the local support we offered may have confused learners because of the many distractors that were provided in addition to the relevant elements to choose from. As a consequence, the support may have involved an assessment of the students' understanding rather than just helping them to construct a meaningful representation. The purpose of including distractors had been to still involve the students in

some kind of generative activity, where they have to decide which elements would be best suited to create an adequate drawing. Thus, they would have to retrieve their knowledge acquired from the text and apply it to the creating of the drawing. At this point, we cannot rule out that this retrieval practice may have been harmful rather than helpful. Accordingly, future studies should contrast the present type of local support with a condition where only the relevant elements are presented; moreover, students might be asked to copy these elements rather than using them to assemble a drawing by means of drag-and-drop. The latter would make the support condition more similar to the free-hand drawing condition in that it also involves creating an image (based on a template).

We want to emphasize that the drawings we investigated in the present study are representational drawings, which visually and spatially represent the real-world properties of those objects (Alesandrini 1984; Carney and Levin 2002). These learner-generated drawings can also be referred to as schematized illustrations and they are intended to depict what is described in the written text. This definition excludes other types of visual representations such as realistic pictures but also nonrepresentational visualizations such as diagrams or graphs because the underlying cognitive processes that learners engage in to generate them diverge (e.g., Ainsworth 1999). Thus, no conclusion can be derived for nonrepresentational drawings based on the present findings.

Conclusion

To sum up, the present study contributes to research on learner-generated drawing. We investigated which types of drawing support are especially helpful for learners to benefit more from constructing drawings during learning. Our findings indicate that neither reducing the requirement to reason about the overall composition of the drawing (global support), nor reducing the need to reason about the visual appearance of single elements (local support) seemed to influence learning beyond an unsupported drawing effect.

To conclude, instructing learners to generate drawings during learning can increase deeper understanding of the learning content irrespective of whether the drawings are constructed free-hand or with additional support. Moreover, learning outcomes depend on the quality of constructed drawings. More specifically, more accurate drawings are associated with higher test performance. Future research should address under which circumstances additional drawing aids can increase the benefits of learner-generated drawing and whether less experienced learners benefit more from support than more experienced learners (Fiorella and Zhang 2018).

Acknowledgements Open Access funding provided by Projekt DEAL. This research was supported as part of the LEAD Graduate School and Research Network [Grant No. GSC1028], a project of the Excellence Initiative of the German federal and state governments.

Funding This study was funded by the Leibniz-Institut für Wissensmedien Tübingen.

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflict of interest.

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

- Ainsworth, S. (1999). The functions of multiple representations. *Computers & Education*, *33*, 131–152. [https://doi.org/10.1016/s0360-1315\(99\)00029-9](https://doi.org/10.1016/s0360-1315(99)00029-9).
- Ainsworth, S., Prain, V., & Tytler, R. (2011). Drawing to learn in science. *Science*, *333*(6046), 1096–1097. <https://doi.org/10.1126/science.1204153>.
- Alesandrini, K. L. (1981). Pictorial-verbal and analytic-holistic learning strategies in science learning. *Journal of Educational Psychology*, *73*, 358–368. <https://doi.org/10.1037/0022-0663.73.3.358>.
- Alesandrini, K. L. (1984). Pictures and adult learning. *Instructional Science*, *13*, 63–77. <https://doi.org/10.1007/BF00051841>.
- Bjork, E. L., & Bjork, R. A. (2011). Making things hard on yourself, but in a good way: Creating desirable difficulties to enhance learning. In M. A. Gernsbacher, R. W. Pew, L. M. Hough, & J. R. Pomerantz (Eds.), *Psychology and the real world: Essays illustrating fundamental contributions to society* (pp. 56–64). New York, NY: Worth Publishers.
- Carney, R. N., & Levin, J. R. (2002). Pictorial illustrations still improve students' learning from text. *Educational Psychology Review*, *14*, 5–26. <https://doi.org/10.1023/A:1013176309260>.
- Chi, M. T. H. (2009). Active-constructive-interactive: A conceptual framework for differentiating learning activities. *Topics in Cognitive Science*, *1*, 73–105. <https://doi.org/10.1111/j.1756-8765.2008.01005.x>.
- Chi, M. T. H., & Wylie, R. (2014). The ICAP framework: Linking cognitive engagement to active learning outcomes. *Educational Psychologist*, *49*, 219–243. <https://doi.org/10.1080/00461520.2014.965823>.
- Cierniak, G., Scheiter, K., & Gerjets, P. (2009). Explaining the split-attention effect: Is the reduction of extraneous cognitive load accompanied by an increase in germane cognitive load? *Computers in Human Behavior*, *25*, 315–324. <https://doi.org/10.1016/j.chb.2008.12.020>.
- Ekstrom, R. B., French, J. W., & Harman, H. H. (1976). *Manual for kit of factor-referenced cognitive tests*. Princeton, NJ: Educational Testing Service.
- 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>.
- Fiorella, L., & Mayer, R. E. (2015). *Learning as a generative activity: Eight learning strategies that promote understanding*. New York: Cambridge University Press.
- Fiorella, L., & Zhang, Q. (2018). Drawing boundary conditions for learning by drawing. *Educational Psychology Review*, *30*, 1115–1137. <https://doi.org/10.1007/s10648-018-9444-8>.
- Greene, T. R. (1989). Children's understanding of class inclusion hierarchies: The relationship between external representation and task performance. *Journal of Experimental Child Psychology*, *48*, 62–89. [https://doi.org/10.1016/0022-0965\(89\)90041-6](https://doi.org/10.1016/0022-0965(89)90041-6).
- Johnson-Laird, P. N. (2005). Mental models and thought. In K. J. Holyoak & R. G. Morrison (Eds.), *The cambridge handbook of thinking and reasoning* (pp. 185–208). Cambridge, NY: Cambridge University Press.
- Koedinger, K. R., & Alevan, V. (2007). Exploring the assistance dilemma in experiments with cognitive tutors. *Educational Psychology Review*, *19*, 239–264. <https://doi.org/10.1007/s10648-007-9049-0>.
- Kosslyn, S. M. (1994). *Image and brain: The resolution of the imagery debate*. Cambridge, MA, US: The MIT Press.
- Lansing, K. M. (1984). The effect of drawing on the development of mental representations: A continuing study. *Studies in Art Education*, *25*, 167–175.
- Leopold, C., & Leutner, D. (2012). Science text comprehension: Drawing, main idea selection, and summarizing as learning strategies. *Learning and Instruction*, *22*, 16–26. <https://doi.org/10.1016/j.learninstruc.2011.05.005>.
- Leopold, C., Sumfleth, E., & Leutner, D. (2013). Learning with summaries: Effects of representation mode and type of learning activity on comprehension and transfer. *Learning and Instruction*, *27*, 40–49. <https://doi.org/10.1016/j.learninstruc.2013.02.003>.

- Lesgold, A. M., DeGood, H., & Levin, J. R. (1977). Pictures and young children's prose learning: A supplementary report. *Journal of Reading Behavior*, 9, 353–360. <https://doi.org/10.1080/10862967709547240>.
- Lesgold, A. M., Shimron, J., Levin, J. R., & Guttman, J. (1975). Pictures and young children's learning from oral prose. *Journal of Educational Psychology*, 67, 636–642. <https://doi.org/10.1037/0022-0663.67.5.636>.
- Leutner, D., Leopold, C., & Sumfleth, E. (2009). Cognitive load and science text comprehension: Effects of drawing and mentally imagining text content. *Computers in Human Behavior*, 25, 284–289. <https://doi.org/10.1016/j.chb.2008.12.010>.
- Leutner, D., & Schmeck, A. (2014). The generative drawing principle in multimedia learning. In R. E. Mayer (Ed.), *The cambridge handbook of multimedia learning* (2nd ed., pp. 433–448). Cambridge, NY: Cambridge University Press.
- Liepmann, D., Beauducel, A., Brocke, B., & Amthauer, R. (2007). *Intelligenz-Struktur-Test 2000 R* (2nd ed.). Göttingen: Hogrefe.
- Mason, L., Lowe, R., & Tornatora, M. C. (2013). Self-generated drawings for supporting comprehension of a complex animation. *Contemporary Educational Psychology*, 38, 211–224. <https://doi.org/10.1016/j.cedpsych.2013.04.001>.
- Mayer, R. E. (2002). Multimedia learning. *Psychology of Learning and Motivation*, 41, 85–139. [https://doi.org/10.1016/S0079-7421\(02\)80005-6](https://doi.org/10.1016/S0079-7421(02)80005-6).
- Mayer, R. E. (2014). The cognitive theory of multimedia learning. In R. E. Mayer (Ed.), *The cambridge handbook of multimedia learning* (2nd ed., pp. 43–71). Cambridge, NY: Cambridge University Press.
- Paas, F. G. W. C. (1992). Training strategies for attaining transfer of problem-solving skill in statistics: A cognitive-load approach. *Journal of Educational Psychology*, 84, 429–434. <https://doi.org/10.1037/0022-0663.84.4.429>.
- Paas, F., Renkl, A., & Sweller, J. (2003). Cognitive load theory and instructional design: Recent developments. *Educational Psychologist*, 38, 1–4. https://doi.org/10.1207/S15326985EP3801_1.
- Scaife, M., & Rogers, Y. (1996). External cognition: How do graphical representations work? *International Journal of Human-Computer Studies*, 45, 185–213. <https://doi.org/10.1006/ijhc.1996.0048>.
- Scheiter, K., Schleichok, K., & Ainsworth, S. (2017). Why sketching may aid learning from science texts: Contrasting sketching with written explanations. *Topics in Cognitive Science*. <https://doi.org/10.1111/tops.12261>.
- Schmeck, A., Mayer, R. E., Opfermann, M., Pfeiffer, V., & Leutner, D. (2014). Drawing pictures during learning from scientific text: Testing the generative drawing effect and the prognostic drawing effect. *Contemporary Educational Psychology*, 39, 275–286. <https://doi.org/10.1016/j.cedpsych.2014.07.003>.
- Schmidgall, S. P., Eitel, A., & Scheiter, K. (2019). Why do learners who draw perform well? Investigating the role of visualization, generation and externalization in learner-generated drawing. *Learning and Instruction*, 60, 138–153. <https://doi.org/10.1016/j.learninstruc.2018.01.006>.
- Schnotz, W. (2014). Integrated model of text and picture comprehension. In R. E. Mayer (Ed.), *The Cambridge handbook of multimedia learning* (2nd ed., pp. 72–103). Cambridge, NY: Cambridge University Press.
- Schwaborn, A., Mayer, R. E., Thillmann, H., Leopold, C., & Leutner, D. (2010). Drawing as a generative activity and drawing as a prognostic activity. *Journal of Educational Psychology*, 102, 872–879. <https://doi.org/10.1037/a0019640>.
- Schwaborn, A., Thillmann, H., Opfermann, M., & Leutner, D. (2011). Cognitive load and instructionally supported learning with provided and learner-generated visualizations. *Computers in Human Behavior*, 27, 89–93. <https://doi.org/10.1016/j.chb.2010.05.028>.
- Snowman, J., & Cunningham, D. J. (1975). A comparison of pictorial and written adjunct aids in learning from text. *Journal of Educational Psychology*, 67, 307–311. <https://doi.org/10.1037/h0076934>.
- Sweller, J. (2005). Implications of cognitive load theory for multimedia learning. In R. E. Mayer (Ed.), *The cambridge handbook of multimedia learning* (pp. 19–30). New York: Cambridge University Press.
- Van Dijk, T. A., & Kintsch, W. (1983). *Strategies of discourse comprehension*. New York: Academic Press.
- Van Meter, P. (2001). Drawing construction as a strategy for learning from text. *Journal of Educational Psychology*, 93, 129–140. <https://doi.org/10.1037/0022-0663.93.1.129>.
- Van Meter, P., Aleksic, M., Schwartz, A., & Garner, J. (2006). Learner-generated drawing as a strategy for learning from content area text. *Contemporary Educational Psychology*, 31, 142–166. <https://doi.org/10.1016/j.cedpsych.2005.04.001>.
- Van Meter, P., & Firetto, C. M. (2013). Cognitive model of drawing construction. Learning through the construction of drawings. In G. J. Schraw, M. T. McCrudden, & D. R. Robinson (Eds.), *Learning through visual displays* (pp. 247–380). Charlotte, N.C.: Information Age Publishing.

Van Meter, P., & Garner, J. (2005). The promise and practice of learner-generated drawing: Literature review and synthesis. *Educational Psychology Review*, 17, 285–325. <https://doi.org/10.1007/s10648-005-8136-3>.

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