# Drawing to Learn: Investigating the Role of Contributing Factors and Instructional Support for Learner-Generated Drawing

**Dissertation**

der Mathematisch-Naturwissenschaftlichen Fakultät
der Eberhard Karls Universität Tübingen
zur Erlangung des Grades eines
Doktors der Naturwissenschaften
(Dr. rer. nat.)

vorgelegt von
Dipl.-Psych. Steffen Patrick Schmidgall
aus Tübingen

Tübingen
2017

Gedruckt mit Genehmigung der Mathematisch-Naturwissenschaftlichen Fakultät der Eberhard Karls Universität Tübingen.

| | |
|---|---|
| Tag der mündlichen Qualifikation: | 21.07.2017 |
| Dekan: | Prof. Dr. Wolfgang Rosenstiel |
| 1. Berichterstatter: | Prof. Dr. Katharina Scheiter |
| 2. Berichterstatter: | Prof. Dr. Caterina Gawrilow |

## Acknowledgments

I want to thank everybody who supported me during the journey towards my PhD by discussing study designs, results, and analyses or by offering advice and feedback. Moreover, I want to thank the members of the IWM – especially my colleagues of the Multiple Representations Lab – and the LEAD Graduate School and Research Network for offering me an outstanding and interdisciplinary research environment. I am very grateful that I could be a part of the team.

My deepest gratitude goes to a number of people without whom I wouldn’t have made it to this day…

- Katharina Scheiter and Alexander Eitel who did a great job in being my supervisors
- Caterina Gawrilow who stepped in last minute as second reviewer
- Stephan Schwan who offered helpful feedback
- All my colleagues of the Multiple Representations Lab who made the hard work of the past years so much fun
    - Juliane Kant, Marie-Christin Krebs, and Katrin Schleinschok for proofreading in times when the letters started to dance before my eyes
- Special thanks go to my friends Marie-Christin Krebs and Katrin Schleinschok who encouraged me to take on some marvelous and life-changing adventures
- The research assistants of the Multiple Representations Lab who helped me with data collection and coding
- My parents for always being there for me
- Phillip Lang-Schmidgall for being as patient and supportive as he is, as well as for his unconditional love and support. Without you, I never would have considered drawing (a sheep).

# Table of Contents

# 1 General Introduction

Visual representations are omnipresent in formal as well as informal educational contexts such as schools or museums. For example, a vast majority of text books that are used in schools include some form of visual representation, ranging from decorative pictures that are used to attract learners' interest in a certain topic, to representative visualizations of physical systems, graphs to visualize data, and so forth. Whereas decorative pictures have been shown to be not helpful for learning, representative visual representations are included based on the assumption that they will help learners to understand the learning content better. However, it is important that learners use external visual representations properly in order to be able to benefit from them. First of all, learners need to pay an adequate amount of attention to provided visual representations. That is, learners need to look at these illustrations and process them. Most importantly, information from the visual representations needs to be integrated with the information from the text accompanying visual representations. Unfortunately, learners tend to ignore visual representations, which is why sometimes providing visual representations does not increase learning. However, visual representations do not necessarily need to be provided by a teacher or an author; they can also be learner-constructed. This is the case, for example, when learners design a concept map of a topic or when they engage in drawing of what is explained verbally.

Drawing to learn has been of interest to educational researchers for several decades, with a peak in the late 1970s and early 1980s. Then, for almost ten years, not much attention was given to that topic. However since the late 1990s, there has been a revival of research on drawing to learn. Constructing a drawing of what is explained in a written text has been shown to promote learning throughout different content domains and age groups. Drawing involves generating a visual representation of a verbally presented content that is constructed mentally first and is then externalized into the real world. This strategy is assumed to promote learning because learners need to thoroughly process visual information in order to be able to construct a drawing. Moreover, they engage in a generative task while constructing a drawing, which is also considered to be helpful for learning. The aim of the present dissertation was twofold: First, I aimed to investigate what factors mainly contribute to the benefits of drawing. Second, since constructing a drawing during learning can be a challenging task, I studied whether and what type of instructional support is most effective to help learners benefit more from drawing.

The present thesis is structured into three main parts. First, Chapter 1 provides the theoretical background, including an introduction to learning with author-generated and learner-generated visualizations, and derives the thesis' research questions. Second, Chapters 2 and 3 report three empirical studies that build the centerpiece of this thesis. Studies 1 and 2 focused on the main contributing factors that play a key role in why a drawing strategy instruction is beneficial for learning. Study 3 investigated how instructional support during drawing may improve drawing to learn. Lastly, Chapter 4 discusses the findings of the three empirical studies and their implications for theory and practice.

## 1.1 Learning with External Pictorial Representations

Learner-generated drawing can be viewed as a subtype of learning with text and pictures (multimedia), which has been a focus of psychological and educational research for the past decades. Within a drawing strategy, learners are not provided with an external pictorial representation but they rather construct it themselves. Thus, they learn from a combination of text and self-constructed pictures, which is somewhat analogous to learning with text and provided pictures except for the generative component required during drawing. Because similar processes are important within learner-generated drawing and multimedia learning, the following paragraphs will focus on two theoretical frameworks for learning with multimedia – the cognitive theory of multimedia learning and the integrated model of text and picture comprehension – and also on the effectiveness and challenges of multimedia learning.

### 1.1.1 Cognitive theory of multimedia learning

The cognitive theory of multimedia learning (CTML; Mayer, 2009, 2014a) is a theoretical framework for learning with text and graphics based on assumptions of the dual coding theory (Paivio, 1986), models of working memory (Baddeley, 1992), the generative theory of learning (Wittrock, 1990), and the cognitive load theory (Sweller, 1999, 2005a). Following the multi-store model of memory (Atkinson & Shiffrin, 1968), the CTML also distinguishes three memory stores within the human processing system including sensory memory, working memory, and long-term memory (Mayer, 2009, 2014a). First, incoming information enters the sensory memory that very briefly holds exact sensory copies of incoming words and pictures. Then, information enters the working memory that allows for manipulating selected incoming information. Lastly, information enters the long-term memory where organized knowledge is permanently stored. A graphical representation of the CTML can be found in Figure 1. There are three central assumptions within the CTML: the dual-channel assumption, the limited-capacity assumption, and the active processing assumption.

The dual channel assumption asserts that learners possess two separate processing channels – one for visual/ spatial and one for auditory/ verbal information. This assumption is closely associated with Paivio's dual coding theory (Clark & Paivio, 1991; Paivio, 1986, 2006) and Baddeley's model of working memory (Baddeley, 1999; Baddeley, Eysenck, & Anderson, 2015). When information is presented to the eyes (e.g., graphics, diagrams, videos or written text) processing of that information starts in the visual channel. When information is presented to the ears (e.g., narrations or nonverbal sounds) processing of that information starts in the verbal channel. The difference between these channels can be conceptualized based on representation modes or sensory modalities. The representation-mode approach concentrates on the way a stimulus is presented: verbally (e.g., spoken or printed words) or nonverbally (e.g., pictures, videos or background sounds). One channel processes verbal material, whereas the other channel processes pictorial material and nonverbal sounds. This approach is most consistent with Paivio's (Paivio, 1986, 2006) distinction between a verbal and a nonverbal system. The sensory-modality approach concentrates on whether learners initially perceive information through their eyes (e.g., pictures, videos, animations, printed words) or ears (e.g., spoken words, background sounds). It assumes that one channel processes visually presented material whereas the other channel processes auditorily presented material. This approach is most in line with Baddeley's (Baddeley, 1999; Baddeley et al., 2015) distinction between the visuospatial sketchpad and the phonological loop. The major distinction between these two approaches regarding multimedia learning rests in the processing of printed words and background sounds. Printed words are processed in the verbal channel in the representation-mode approach but in the visual channel in the sensory-modality approach. Background sounds (also including nonverbal music) are processed in the nonverbal channel in the representation-mode approach, whereas they are processed in the auditory channel in the sensory-modality approach. Although information enters the information system via only one channel, learners can translate a representation for processing into the other channel (cross-channel representations). For example, written words are initially processed in the visual channel because they are presented to the eyes, but can then be converted in the working memory into sounds, which are processed through the auditory channel. Likewise, a picture is initially processed in the visual channel, but learners may be able to mentally construct a corresponding verbal description, which is processed in the auditory channel (Mayer, 2014a).

The second assumption of the CTML is the limited-capacity assumption which is based on Baddeley's (Baddeley, 1999; Baddeley et al., 2015) model of working memory and the cognitive load theory (Sweller, 1999, 2005a). The working memory can only process a limited amount of information at the same time. Thus, learners can process complex materials only sequentially and not at once. These constraints in processing capacity force learners to make decisions about which parts of incoming information they pay attention to. Moreover, the constraints affect the degree to which learners build connections among the selected pieces of information and with their existing prior knowledge. Learners with extensive prior knowledge can process more pieces of information at once, because they can group separate information elements into higher-order chunks (Mayer, 2011; Miller, 1956).

The third assumption of the CTML is the active processing assumption, which is based on Wittrock's generative theory of learning (Wittrock, 1990). Within the CTML, learners are not viewed as passive recipients of information but rather as active processors who seek to make sense of multimedia presentations. Active processing can be viewed as a process of model building because the desired outcome of active cognitive processing is the construction of a coherent mental representation of the to-be-learned content. Such a mental model represents the key elements of the presented material and their relations. For example, in a multimedia presentation of how a pulley system works, learners may attempt to build a cause-and-effect system. Learners can then inspect this system and figure out what a change in one part of the system causes in another part of the system. There are three cognitive processes that are essential for actively building a coherent mental model within the CTML: selecting relevant verbal and pictorial information, organizing selected verbal and pictorial information, and integrating selected verbal and pictorial information with existing prior knowledge (Mayer, 2009, 2014a; Wittrock, 1990; see also Figure 1).

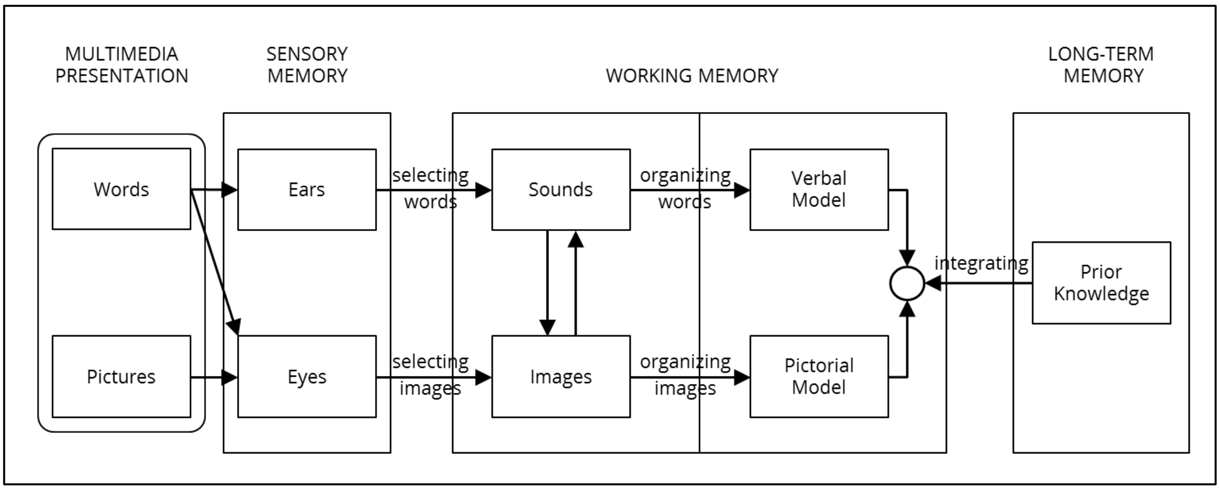


*Figure 1*. Cognitive theory of multimedia learning (Mayer, 2014a, p. 52).

Selecting relevant words and picture elements occurs within the sensory memory when learners pay attention to some of the words and images that are presented within multimedia material. The requirement to select only parts of the presented verbal and pictorial information arises from the limited working memory capacity. It is neither possible for learners to process all presented words, nor all parts of a complex visualization, so they must focus on only part of the incoming information at a time. Moreover, selecting relevant words and picture elements involves bringing information from the outside into the working memory component of the cognitive system. Thus, an initial surface verbal representation (sounds) and an initial surface pictorial representation (images) are built in the working memory. The processing of written text additionally requires learners to transfer the text from the visual system to the verbal system in order to build an auditory representation of the written text. Finally, selection is no arbitrary process because learners need to determine which words and picture elements are most relevant. Therefore, this activity is also in line with the perspective of the learner as an active sense maker rather than a passive consumer (c.f. active processing assumption).

Organizing selected words and picture elements occurs when learners build structural relations among the elements within one surface representation in order to construct a coherent verbal mental model and a coherent pictorial mental model. This process is assumed to take place in the auditory channel for building a coherent verbal model, and is assumed to take place in the visual channel for building a coherent pictorial model. Moreover, this process is also influenced by the limited working memory capacity. Therefore, learners lack the ability to build all possible connections among the elements within each surface representation. As a result, they rather have to focus on building a simple structure that makes sense to the learner such as a cause-and-effect chain.

Integrating the verbal and pictorial models with prior knowledge involves building connections between the verbal mental model and the pictorial mental model while considering relevant portions of learners' existing prior knowledge. This process requires learners to activate knowledge in long-term memory and bringing it into working memory. It involves a vast amount of cognitive resources, so existing prior knowledge can facilitate integration and reduce learners' cognitive load. The aim of integration is that learners build a coherent mental model of the learning content and store it in long-term memory. Moreover, integration is assumed to be the most relevant process involved in learning with multimedia (Bodemer, Ploetzner, Feuerlein, & Spada, 2004; Schüler, Arndt, & Scheiter, 2015; Seufert, 2003).

### 1.1.2 Integrated model of text and picture comprehension

Another framework to understand learning with multimedia material was introduced with the integrated model of text and picture comprehension (ITPC; Schnotz, 2014; Schnotz & Bannert, 2003). The ITPC and the CTML (Mayer, 2009, 2014a) have a great overlap as both frameworks emphasize the importance of selection, organization, and integration processes.

Therefore, the following paragraph will focus mainly on the differences between the two frameworks.

Similar to the CTML, the ITPC assumes that during multimedia learning cognitive processes of selection and organization of information as well as integration of information in consideration of relevant existing prior knowledge takes place. The ITPC also differentiates between two cognitive systems of information processing which are limited in capacity. Whereas the CTML distinguishes between words and pictures with regard to representation mode, the ITPC distinguishes between the text as descriptive representation (containing symbols) and realistic pictures as depictive representation (containing icons). Symbols are arbitrary structures that are connected with the referent through conventions. Icons are analogous representations which are associated with their referent by similarity or another structural commonality. According to Schnotz (2014), the particular benefit of realistic pictures results from this distinction. Because these realistic pictures illustrate the represented object in a structural or functional analogous way, they are based on similar principles of representation as mental models. For example, in learning material about migratory birds a map and a drawing of a bird are pictorial objects that share some similarities with their corresponding referents (i.e., the European and African continents or the white stork). A bar graph within the same learning material has a more abstract structural commonality with its referent. The meaning of the bar graph, for example, can be based on an analogy that the height of the bars correspond to the number of white storks observed in a certain habitat during a certain month, whereas the sequence of bars can correspond to the sequence of months during the year. Thus, a mental model derived from depictive pictures that correspond to their respective referents by means of analogy can be constructed more efficiently than a mental model derived from text.

Moreover, the ITPC characterizes the integration process of the verbal and the pictorial model into a coherent mental model in a more sophisticated way than the CTML. In contrast to the CTML, the mental transformation processes do not occur on the bottom level of representation in working memory when information enters the working memory. According to the ITPC, these integration processes take place on a higher level in working memory between the propositional representation and the mental model. Both subsystems contain information from the descriptive and the depictive channel and interact closely. Understanding of a written text with provided illustrations following the ITPC is explained in the following (see Figure 2).

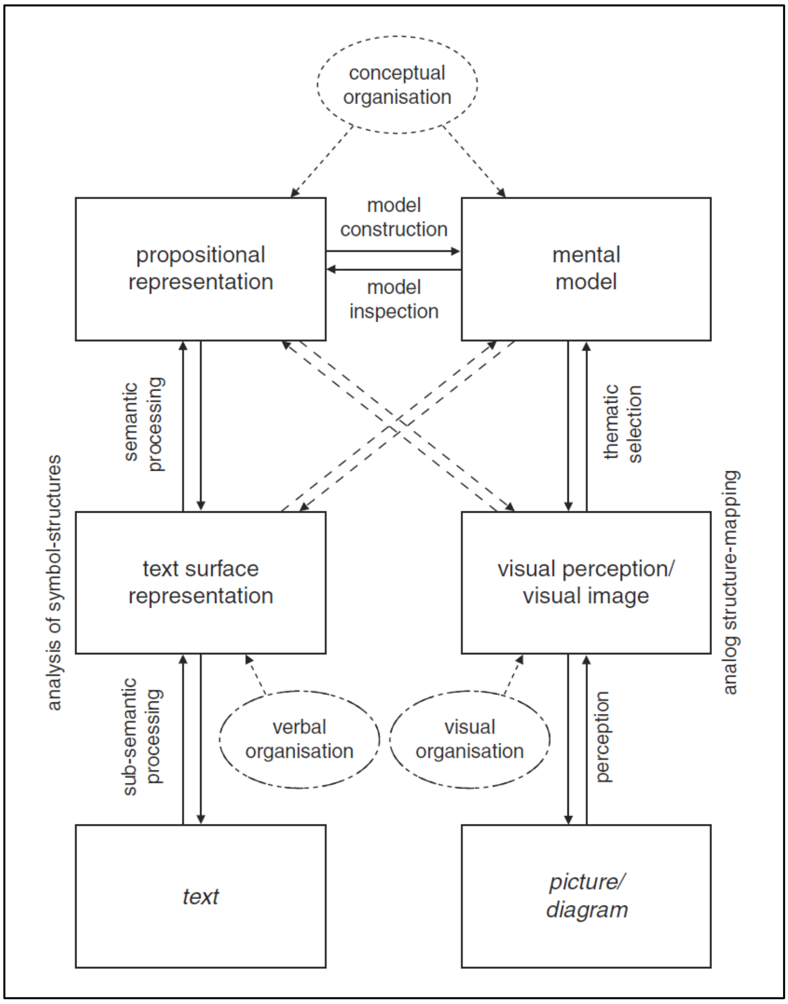


*Figure 2*. Schnotz and Bannert's (2003) integrated model of text and picture comprehension (Schnotz, 2014, p. 79).

First, written text enters the visual register through the eyes and is transmitted to the working memory where a surface representation of the features of the text is created. Through semantic processing, relevant text information is selected and transferred into a propositional representation. Then, the propositional representation is extended to a mental model by taking topic-specific schemes from long-term memory (prior knowledge) and existing representations from the depictive channel into account.

Pictures are initially perceived through the eyes and enter the visual register. Then, they are subjected to visual feature analysis resulting in visuo-spatial patterns in working memory as a visual perceptual representation (visual perception/ visual image). Subsequently, depictive processing through the mapping of selected perceptual structures leads to the construction or elaboration of a corresponding mental model. Moreover, the mental model can be expanded by applying relevant topic-specific schemes from long-term memory as well as relevant information of the propositional representation of a corresponding text. Additionally, this model can be used to extract new information through reading information or drawing inferences that can be encoded in the propositional representation (model inspection).

Thus, the propositional representation can be used to construct a mental model and, on the other hand, the mental model can be used to construct a propositional representation through model inspection. In contrast to the CTML, which assumes that learners first build separate mental representations for verbal and pictorial information (verbal model vs. pictorial model), the ITPC postulates that learners construct a joint representation of text and pictures early on. The CTML emphasizes the integration of the verbal and the pictorial model, whereas the ITPC emphasizes the interaction between propositional representation and mental model.

### 1.1.3 Effectiveness of learning with text and pictures

The multimedia principle refers to the robust research finding that learning from a combination of text and pictures is more effective than learning from text alone (Butcher, 2014). The multimedia principle could be empirically validated in numerous studies using paper-based instructional material (Carney & Levin, 2002; Mayer, 1989; Mayer & Anderson, 1991, 1992; Mayer & Gallini, 1990; Plass, Chun, Mayer, & Leutner, 1998) as well as computer-based learning environments (Butcher, 2006; Cuevas, Fiore, & Oser, 2002; Eitel, 2016; Eitel, Scheiter, & Schüler, 2013; Eitel, Scheiter, Schüler, Nyström, & Holmqvist, 2013; Mayer & Moreno, 2002). Multimedia learning has been shown to be especially relevant in domains whose content covers large proportions of visuo-spatial information like biology (Stalbovs, Scheiter, & Gerjets, 2015), physics (Mayer, 1989; Mayer & Gallini, 1990) or geography (Harp & Mayer, 1997).

However, the body of research on multimedia learning also shows that learners do not necessarily benefit from these instructional materials. In eye-tracking experiments, in which the movements of learners' eyes are recorded during learning, it has been shown that learners tend to ignore the pictures while learning; thus, showing an overreliance on text (Hegarty & Just, 1993). Moreover, learners show only few attempts to integrate verbal and pictorial information (Cook, Wiebe, & Carter, 2008; Mason, Tornatora, & Pluchino, 2013; Scheiter & Eitel, 2015).

Mason et al. (2013) were interested in whether young children already show attempts to integrate verbal and pictorial information while constructing a coherent model from the two respective types of representation during reading. Fourth-graders were presented with a multimedia lesson about the topic of air during which their eye movements were recorded. Eye-tracking revealed three different patterns of processing the materials (high vs. intermediate vs. low integrators) and it was integration between verbal and pictorial information that differentiated these patterns essentially. Moreover, higher integration behavior was related to higher learning outcomes emphasizing the importance of integrating the verbal with the pictorial representation while learning from an illustrated text.

A coherent mental model of the learning content that integrates verbal and pictorial information is important for multimedia learning (Mayer, 2009, 2014a). While integrating, learners build one-to-one connections between the verbal and pictorial mental model. In order to do that, learners must map elements, actions, and causal relations in the verbal representation to elements, actions, and causal relations in the pictorial model (Mayer, 1997; Schüler, 2017). Both, the CTML and the ITPC, postulate that connecting verbal and pictorial information to one another is the most central step in learning with text and pictures, because integration does not only support remembering information, but also fosters deeper understanding of the multimedia message's content (Mayer, 2009, 2014a; Schnotz, 2014; Schnotz & Bannert, 2003). Research has shown that learners are able to relate verbal information to pictorial information when being asked to recall information (Glenberg & Langston, 1992; McNamara, Halpin, & Hardy, 1992). Moreover, it has been shown that text-picture-integration occurs already during learning in that successful learners switch more frequently between text processing and picture processing (Hannus & Hyönä, 1999; Hegarty & Just, 1993; Mason, Tornatora, & Pluchino, 2015; Mason et al., 2013; O'Keefe, Letourneau, Homer, Schwartz, & Plass, 2014). Because learners show an over-reliance on text, tend to ignore provided pictures, and thus show only few attempts to integrate, it has been of interest how to foster text-picture-integration.

In two eye-tracking experiments, Scheiter and Eitel (2015) investigated how highlighting text-picture correspondences (so called signals) affect visual attention and learning outcomes in a multimedia message. Students received multimedia material about the circulatory heart system either with or without signals that highlighted the correspondences between important parts of the text in the provided visualizations. The results showed an advantage of signals reflected by better learning outcomes of the respective group compared to a control group without signals. The positive effect of signals could be explained by changes in students' visual attention, that is, students attended to signaled information more frequently and earlier during learning than to non-signaled information. Signals positively influenced text-picture integration performance resulting in higher learning outcomes. Highlighting text-picture correspondences is one way to foster integration of verbal and pictorial information among learners in multimedia learning (for an overview see Richter, Scheiter, & Eitel, 2016).

Although learners have both types of representations available in material consisting of an illustrated text, they tend to have difficulties in integrating both types of representations in order to be able to really benefit from multimedia learning. The effectiveness of multimedia learning has been shown to be mediated by learners' integration behavior. Learners who show more attempts to integrate (Mason et al., 2013) and attend to important corresponding information in text and visualizations early on (Scheiter & Eitel, 2015) are more successful than learners who don't show this behavior. Adapting the instructional material in terms of highlighting corresponding important elements of the text in the pictures can promote integration processes. Yet, learners have to rely on the author of the instructional material to optimize it for learning. So another way to address learners' lack of integration attempts, and to guide learners' attention to as well as promote processing of visual information is to instruct learners to self-generate the pictorial representations accompanying a written text. With such a drawing strategy "integration itself is forced as the verbal representation is the foundation for the nonverbal representation" (van Meter & Garner, 2005, p. 318). This strategy, which is known in the literature as learner-generated drawing, is the main focus of the present thesis.

## 1.2 Learner-Generated Drawing: Introduction and Theoretical Framework

Learner-generated drawing requires learners to construct external pictorial representations that include the key concepts and their relations while learning from verbal instruction (Leutner & Schmeck, 2014). It is important to differentiate between two characteristics of the drawing strategy. First, the drawing is representational which means that the drawing resembles the real-world properties of the depicted relevant objects as well as their spatial relations to one another (Alesandrini, 1984; Carney & Levin, 2002; van Meter & Firetto, 2013). This characteristic excludes nonrepresentational drawings, such as graphs, diagrams, or matrix notes (Ainsworth, 1999; Grossen & Carnine, 1990; Jairam & Kiewra, 2010), because the underlying cognitive processes that learners use to generate them diverge.

Second, the drawings are learner-generated meaning that the learner is the primary causal agent in construction and appearance of the pictorial representation (van Meter & Garner, 2005). Thus, the learner is responsible for determining what the final drawing will look like. This means that the learner is not a passive consumer of information but rather is actively involved in the selection, organization, and integration processes during learning.

The drawing strategy may be as straightforward as paper-pencil drawings (Gobert & Clement, 1999; van Meter, Aleksic, Schwartz, & Garner, 2006), may involve more elaborated colored drawings (Scheiter, Schleinschok, & Ainsworth, in press) or may also include additional drawing aids like cut-out figures of story characters to construct pictorial representations of story events (Lesgold, Shimron, Levin, & Guttmann, 1975). Additional drawing aids may also include a legend that depicts the visual appearance of several structures that are relevant for drawing construction (Schmeck, Mayer, Opfermann, Pfeiffer, & Leutner, 2014).

### 1.2.1 The cognitive model of drawing construction

The following paragraphs will focus on a comprehensive theoretical framework of learner-generated drawing: The cognitive model of drawing construction (CMDC; van Meter & Firetto, 2013). It is an extension of van Meter and Garner's (2005) generative theory of drawing construction (GTDC) that describes the underlying cognitive processes of drawing. The GTDC was developed by applying Mayer's (2009, 2014a) theoretical assumptions of multimedia learning to drawing highlighting the central processes of selecting, organizing, and integrating information. The CMDC takes perspectives of additional theoretical frameworks into account, namely the ITPC (Schnotz, 2014; Schnotz & Bannert, 2003) and self-regulated learning (Winne & Hadwin, 1998). A graphic depiction of the CMDC is shown in Figure 3.

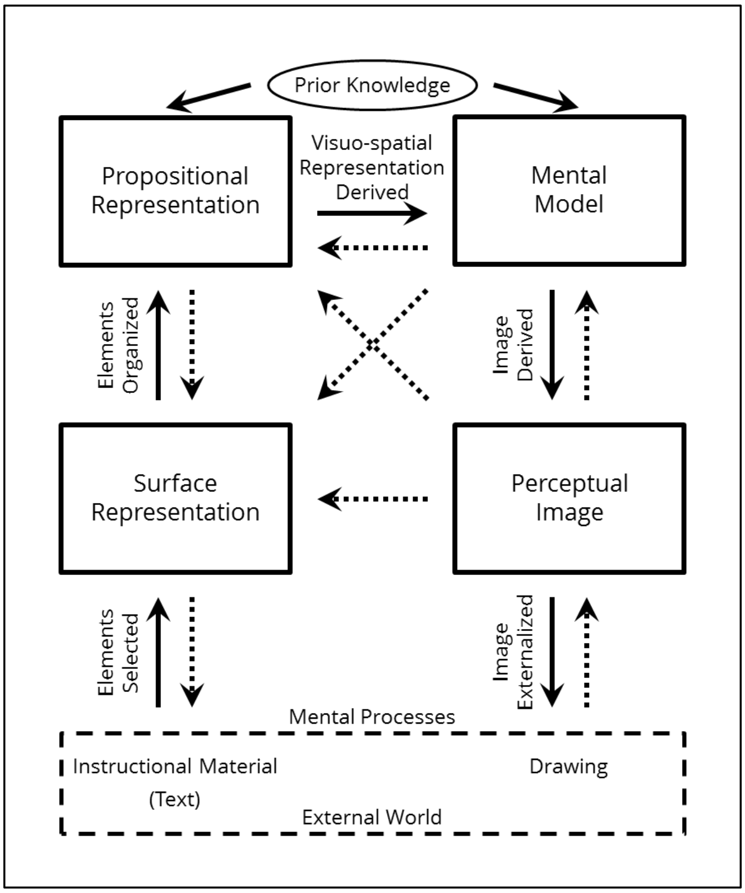


*Figure 3*. Adapted version of the cognitive model of drawing construction (van Meter & Firetto, 2013, p. 256).

The boxes with solid lines in this figure indicate learners' mental knowledge representation, whereas the box with dashed lines at the bottom indicates the representations present in the external world. The labels of the mental representations have been taken from the ITPC to differentiate between descriptive and depictive representations (Schnotz, 2014; Schnotz & Bannert, 2003). The arrows with solid lines reference the cognitive processes that learners take from the provided verbal instructional material to the externalized self-generated drawing.

According to the CMDC, drawing starts with the verbal instructional material. To begin with, learners form a mental surface representation of linguistic features of the text during initial reading. In this first step, learners select the key elements that are most important for the learning content from which to build the surface representation. By organizing the selected elements through semantic processing, a propositional network is constructed that describes structural elements and their relations. Based on the propositional network, learners build a mental model to produce a drawing by deriving a visuo-spatial representation. Consequently, the propositional network defines which structures are included in the mental model, their external appearance, and how they are related to one another. In line with the ITPC, van Meter and Firetto (2013) assume that the mental model is depictive, may include visuo-spatial information, and is more determinant than the propositional network (Cox, 1999; Gobert & Clement, 1999). Thus, the mental model integrates semantic with visuo-spatial information and represents structural relations in a way that allows learners to understand system components and their joint operation like a cause-and-effect chain. The mental model is assumed to be primarily responsible for beneficial effects of drawing on learning (van Meter & Firetto, 2013).

Prior knowledge influences both the propositional network and the mental model which is indicated by the oval in the depiction of the CMDC. It plays an especially important role when learners construct a drawing in the absence of any provided pictorial representation. For example, when trying to translate a part of the text reading “The left exterior is convex”, learners must consult their long-term memory to determine how the word “convex” can be translated into pictorial form.

Based on the mental model, learners derive a perceptual image in order to construct a drawing. The perceptual image is a depictive surface feature representation that learners can externalize onto paper as a drawing. Although only a single arrow is entering the perceptual image, van Meter & Firetto (2013) emphasize that also, for example, aspects of the surface representation can be translated directly to the perceptual image but that the drawing strategy is most effective when the perceptual image is derived from the mental model.

The dashed arrows in the depiction of the CMDC emphasize the role of metacognition. These arrows indicate feedback cycles that are triggered when learners notice that content is not well understood while attempting to draw. For example, learners may need clarification regarding the relation between two components during the construction of the mental model and reconsider the propositional network. Also, efforts to derive the perceptual image may lead learners back to inspecting the surface representation to identify relevant information. These feedback cycles may even lead learners to the realization that reinspection of the instructional material is useful. Thus, drawing is not viewed as a simple linear sequence but rather as an iterative process driving learners back and forth between the various mental and external representations.

Moreover, the CMDC incorporates principles of self-regulated learning involving setting standards for performance, applying operations, and monitoring goal progress (Winne & Hadwin, 1998). First, when learners are instructed to “make a drawing”, goal standards are being set. These standards reflect the learners’ understanding of how to construct a drawing and their performance standards. Besides, they will anchor metacognitive control while the strategy is executed. Second, cognitive operations are carried out to comprehend the instructional material and to match task standards. These are the operations used to select, organize, and integrate information during drawing. Drawing may also prompt the use of other known learning strategies, for example self-questioning to construct the propositional network. Third, learners monitor their progress towards the drawing goal. Here, metacognitive control is triggered when learners compare their drawing in progress to their standards set for task performance. When the drawing meets the standards, learners will most likely continue to draw the remaining parts to complete the drawing task if the drawing is not yet completed. Otherwise – if the drawing is already completed –, learners will most likely terminate their drawing process. When the standards have not been met, metacognitive control will direct learners back to the mental model, propositional network or even the instructional material. In the present thesis, however, the main focus lies on cognitive processes involved in drawing.

## 1.3 Current Issues in Drawing Research

The following paragraphs will give a brief overview on previously conducted research on drawing illustrating the various aspects of learning that can be positively affected by a drawing strategy.

Constructing a drawing can improve learners' imagination and observational processes because it helps to recognize the details and subtle properties of different objects which is especially important for learning scientific content (Dempsey & Betz, 2001; Steele, 1991; Stein & Power, 1996). Drawing can also increase content area knowledge because drawing engages learners in higher-level thinking and supports a deep understand of the material (Britton & Wandersee, 1997; Costantino, 1986; Johnson, 1988; Schwamborn, Thillmann, Opfermann, & Leutner, 2011; Stein & Power, 1996; van Meter, 2001; van Meter et al., 2006). Furthermore, text comprehension can be positively influenced by drawing because it activates background knowledge, supports the use of other strategies and increases the knowledge acquired from text (Fisher, 1976; Hall, Bailey, & Tillman, 1997; Leopold & Leutner, 2012; McConnell, 1993; Rich & Blake, 1994; Schmeck et al., 2014).

In two experiments, Leopold and Leutner (2012) investigated the helpfulness of drawing compared to text-focused learning strategies (main idea selection or summarization). The study followed a 2x2-design with drawing instruction (yes versus no) and text-focused instruction (yes versus no). Students from grade 10 were instructed to either construct drawings or write down a list of the most important concepts/ short summary after reading a science text about water molecules. In the 'both-strategies' group, the drawing strategy and text-focused strategy were combined. A non-strategy control group (neither drawing nor text-focused strategy instruction) provided a baseline for comparison. In both experiments, drawing strategy instruction fostered text comprehension while text-focused strategy instruction hindered learning.

Another benefit of drawing is that it facilitates writing processes as it helps students to plan their writing activities, generate ideas, and aids thought organization (Caldwell & Moore, 1991; Dietz, 1976; Hubbard, 1987; Karnowski, 1986; Moore & Caldwell, 1993). Learners' affective processing may also be positively influenced by drawing, resulting, for example, in stimulated interest or increased levels of involvement and motivation (Costantino, 1986; Fisher, 1976; Johnson, 1988; McConnell, 1993; Moore & Caldwell, 1993).

## 1.4 Factors Contributing to the Drawing Strategy's Benefits

There are at least three factors that may contribute to the positive effects of drawing as a learning strategy. First, learners may benefit from drawing because they are actively engaged in the leaning process. According to the generation effect (Foos, Mora, & Tkacz, 1994; Slamecka & Graf, 1978), learners achieve superior performance for information that they generated themselves. This information may go beyond the provided information and can contain elaborations and ideas that are not explicitly stated in the text (Chi, 2009; Chi & Wylie, 2014). Second, drawing results in an internal visual representation in addition to the verbal code constructed from text (Paivio, 1986). According to the dual coding theory, learning strategies that involve visualization are expected to be more effective than verbal strategies that result in only a linguistic code because the former yield two codes available in memory rather than just one (Paivio, 1991; van Meter & Garner, 2005). Third, during drawing, the visual representation is not only constructed internally, but is externalized, thereby resulting in a multimedia representation. According to the multimedia effect, learners learn better with multimedia than with text alone (Butcher, 2014; Mayer, 2009).

To disentangle these different contributing factors, it is helpful to compare drawing with other learning strategies that rely on only one or two of these factors (see Table 1 for an overview). First, there are strategies that, like drawing, focus on generation. For instance, writing summaries, similar to drawing, involves the generation of information that goes beyond the written text and results in an external representation, but it does not involve a visual code as does drawing. Research on summarization has shown positive effects on comprehension as it supports learners in identifying the main ideas of a text and in building relations between them (Bean & Steenwyk, 1984; Wittrock & Alesandrini, 1990).

Scheiter et al. (2017) investigated whether benefits of drawing can be explained better by generation of additional content during learning or by the construction of an additional pictorial representation. To this end, seventh-graders were either instructed to generate drawings or to generate written self-explanations of an expository text about the greenhouse effect. The results showed that the groups only differed marginally on an overall level of learning outcome performance in that the drawing group tended to score higher than the self-explanation group. A more in depth analysis revealed that while the groups did not differ for students who applied either strategy poorly, the drawing group outperformed the self-explanation group for students who implemented the strategies with higher quality. This finding indicates that benefits of drawing seem to be specifically connected to the factor of visualization and are not only due to generative learning. Unfortunately, no reading-only control condition was included in the design so it remains an open question how their results relate to that baseline.

Table 1

*Contributing factors generation (corresponding to the generation effect), visualization (corresponding to dual coding theory), and externalization (corresponding to the multimedia effect) and their involvement in a number of learning strategies investigated within the present dissertation*

| Learning strategy | Contributing factor | | |
|---|---|---|---|
| | Generation | Visualization | Externalization |
| Drawing | yes | yes | yes |
| Summarizing | yes | no | yes |
| Mental Imagery | yes | yes | no |
| Multimedia | no | yes | yes |
| Observation | no | yes | yes |
| Text-only | no | no | no |

Moreover, asking learners to construct mental images of the structures and processes described in a written text fosters deeper learning (Dunlosky, Rawson, Marsh, Nathan, & Willingham, 2013; Leopold & Mayer, 2014). Mental imagery, similar to drawing, involves the generation of new information represented in a visual code. In order to do that, learners transfer textual information into pictorial information, and thus, construct referential connections between structures of the text and structures of the picture (van Meter & Garner, 2005). Different from drawing, mental imagery does not result in an external representation; the self-constructed visual code stays purely mental.

Second, there are strategies that, unlike drawing, do not involve generation. Different studies have shown that multimedia learning is superior to learning with text alone (cf., Ainsworth, 1999; Butcher, 2014). Similar to drawing, multimedia learning involves processing information in a linguistic and a pictorial code that is externalized; different from drawing, multimedia learning does not involve generation. Analogous to multimedia learning, observing the pictorial representation of a multimedia presentation evolving step by step similar to drawing involves both dual coding of the material in a linguistic and an additional pictorial code as well as externalized representations of the learning content; different from drawing it does not involve generation.

Additionally, the importance of generation for drawing can be explained within the ICAP framework that emphasizes the role of cognitive engagement during learning (Chi & Wylie, 2014). The ICAP framework assigns learning activities to one of four postulated modes that differ with respect to learners' cognitive engagement and that can be defined on the basis of learners' overt behaviors (engagement behaviors): Interactive, Constructive, Active, Passive. In the passive mode, learners receive information from the instructional material as they, for example, read a text. In the active mode, some motoric action or physical manipulation is undertaken, for example, when learners are repeating, underlining or highlighting during reading a text. In the constructive mode, learners generate or produce additional externalized outputs, for example, when learners draw concept maps or take notes using their own words. In the interactive mode, a group of two or more learners engages in meaningful (constructive) exchange about the instructional material that provides feedback, for example, when a group of learners engages in asking and answering of content related comprehension questions. The ICAP hypothesis predicts that when learners become more engaged with the instructional material from passive to active to constructive to interactive higher cognitive engagement is reflected in increased learning. Based on the assumptions of the CTML (Mayer, 2009, 2014a) and the ITPC (Schnotz, 2014; Schnotz & Bannert, 2003), one can argue that multimedia learning represents learning in the active mode. Learner-generated drawing, on the other hand, can be considered a generative learning activity reflecting learning in the constructive mode. Thus, the processes involved in multimedia learning and drawing may also differ with regard to learners' cognitive engagement during learning. Based on the ICAP framework, cognitive engagement should be higher during drawing than during multimedia learning.

Last, learning from an expository text alone does not rely on visualization, generation or externalization.

To disentangle these contributing factors, a set of experiments (Study 1 and Study 2) was conducted comparing drawing to other learning strategies that differ with respect to generation, visualization, and externalization.

## 1.5 Boundary Conditions for Learner-Generated Drawing

Even though drawing is assumed to promote learning as a pathway to mental model construction in a variety of content domains and age groups, research on drawing has produced somewhat mixed results (see Ainsworth, Prain, & Tytler, 2011; Alesandrini, 1984; Leutner & Schmeck, 2014; van Meter & Firetto, 2013; van Meter & Garner, 2005 for overviews) in which some studies reported positive effects on learning outcomes (Alesandrini, 1981; Hall et al., 1997; Leopold & Leutner, 2012; Lesgold, DeGood, & Levin, 1977; Lesgold et al., 1975; Scheiter et al., 2017; Schmeck et al., 2014; Schwamborn, Mayer, Thillmann, Leopold, & Leutner, 2010; van Meter, 2001; van Meter et al., 2006) whereas others did not (Leopold, Sumfleth, & Leutner, 2013; Leutner, Leopold, & Sumfleth, 2009; Rasco, Tennyson, & Boutwell, 1975; Snowman & Cunningham, 1975; Tirre, Manelis, & Leicht, 1979).

Research on learner-generated drawing identified a number of boundary conditions that influence the effectiveness of a drawing strategy instruction. Benefits of drawing seem to depend on the quality of the drawings, the type of posttest used to assess learning, and additional drawing aids (for an overview see van Meter & Firetto, 2013). The following paragraphs will outline previously investigated boundary conditions and introduce test delay as an additional boundary condition that could influence the benefits of drawing.

### 1.5.1 Drawing quality

Benefits of drawing seem to be related to the quality of learners' drawings constructed during learning: Learners who generate high-quality drawings tend to perform better on learning outcome tests than learners who produce low-quality drawings (Greene, 1989; Hall et al., 1997; Lesgold et al., 1975; Leutner et al., 2009; Mason, Lowe, & Tornatora, 2013; Scheiter et al., 2017; Schmeck et al., 2014; van Meter, 2001; van Meter et al., 2006). Here, quality refers to the amount of key elements correctly incorporated into the drawing from the instructional text (e.g., Schmeck et al., 2014). Accordingly, a high-quality drawing is a drawing that is complete with regard to the most relevant elements and their relations stated in the text. The finding that drawing quality is positively associated with learning outcomes can be described as the prognostic drawing effect (Schwamborn et al., 2010, 2011).

### 1.5.2 Type of posttest

Another boundary condition to learner-generated drawing is that positive effects of drawing on learning appear to depend on the type of test used to assess learning outcomes (van Meter & Firetto, 2013; van Meter & Garner, 2005). Beneficial effects of drawing are more likely to arise on assessments of higher-order knowledge than on assessments of lower-order knowledge such as (verbal) recall (van Meter & Firetto, 2013). Assessments of higher-order knowledge for which drawing effects were observed contain, for example, tests that measure (mathematical) problem solving (Arcavi, 2003; de Bock, Verschaffel, & Janssens, 1998; de

Bock, Verschaffel, Janssens, van Dooren, & Claes, 2003; Hembree, 1992; Larkin & Simon, 1987; Rellensmann, Schukajlow, & Leopold, 2016; van Essen & Hamaker, 1990; van Meter, 2001; van Meter et al., 2006), comprehension (Alesandrini, 1981; Leopold & Leutner, 2015), transfer (Leopold & Leutner, 2012; Schwamborn et al., 2010, 2011) or visuo-spatial knowledge in which learners have to draw diagrams depicting key concepts (Lansing, 1984; Leopold & Leutner, 2012; Schmeck et al., 2014; Schwamborn et al., 2011). On the other hand, no drawing effects were observed for assessments of lower-order knowledge including recognition (van Meter, 2001; van Meter et al., 2006) or factual knowledge (Leutner et al., 2009; Snowman & Cunningham, 1975).

Taking the CMDC into account, it seems that beneficial effects of drawing emerge if the test to assess learning outcomes matches the characteristics of the mental model that is constructed through the drawing activity (van Meter & Firetto, 2013). This dependence may also have implications for the prognostic drawing effect. As a consequence, positive correlations between the quality of constructed drawings and learning outcomes should also depend on the type of test. In particular, correlations for tests that assess higher-order knowledge are expected to be higher than correlations for tests that assess lower-order knowledge.

Van Meter (2001) investigated whether constructing a drawing can improve learning from a biology science text. Four groups of five- and six-graders read a text about the central nervous systems and received different instructions. One group received pure drawing instructions, whereas a second group constructed drawings and compared them with provided illustrations. A third group additionally answered prompting questions to guide the comparison between self-generated drawings and provided illustrations. In a forth group, the control condition, students inspected provided illustrations. The results showed positive effects of drawing for the verbal recall measure (higher-order knowledge) but not for the recognition measure (lower-order knowledge). Moreover, pure drawing activity was as beneficial as inspecting provided illustrations after reading a text. Only students who received provided illustrations along with prompting questions outperformed the illustrations-only group. Unfortunately in this experiment, no text-only control condition was included that received only reading instructions and no provided illustrations to provide a baseline for the effectiveness of pure drawing activity.

Van Meter’s (2001) finding that drawing instructions were beneficial only in combination with inspecting provided pictures after drawing along with answering prompting questions nicely points to another before mentioned characteristic: Instructionally provided support that assists learners in the drawing process.

### 1.5.3 Instructional support

Positive effects of drawing seem to depend on the availability and type of support whose function is to constrain and structure the drawing activity (van Meter & Firetto, 2013; van Meter & Garner, 2005). In particular, drawing is more effective when learners’ construction of drawings is assisted by some kind of additional information (Lesgold et al., 1977, 1975; Schmeck et al., 2014; Schwamborn et al., 2010; van Meter, 2001; van Meter et al., 2006). For example, van Meter (2001) and van Meter et al. (2006) provided students with author-generated illustrations after drawing and showed that this approach enhanced the benefits of drawing. The results showed that that students who received illustrations after drawing outperformed students who drew without support. By comparing their own drawings with provided illustrations students found out what their drawings were supposed to look like (and could identify possible misconceptions) which might have lead students back to revise their drawings and, thus, their mental model. According to the CMDC (van Meter & Firetto, 2013), this behavior should improve comprehension.

Even though the provision of illustrations enhances drawing quality by providing feedback, learners might rely too heavily on this feedback. It bares the risk that learners merely copy the provided illustrations, consequently suppressing constructive learning processes of drawing. Moreover, while this intervention was found to be effective, it is important to note that students were supported after an initial drawing was constructed and not directly during the construction process. In the present theses, I wanted to focus on instructionally provided support that is administered directly during the construction process when a learner reads the text and tries to incorporate all important information into a drawing.

Schmeck et al. (2014) provided support during the construction process by presenting students a drawing prompt that included a legend showing all the relevant elements for the drawings and a partially pre-drawn background for constructing drawings. The presented elements could be used as prototypes by learners for their own drawings. Learners who generated a drawing while being scaffolded performed better in a subsequent knowledge test than students who only read the text. In this kind of intervention, the proportion of learners' personal contribution in constructing the shape of relevant main objects of the to be drawn content is reduced, because learners can copy the shape of relevant objects and merely need to figure out how these objects fit together (i.e., their visuo-spatial relations). Furthermore, when provided with a pre-drawn background, learners do not have to focus on the appearance and spatial relations of irrelevant objects and can fully concentrate on the most relevant parts of the drawing. The personal contribution is reduced even further when learners are asked to only select and assemble relevant objects to construct a drawing and do not have to engage in the process of actual drawing at all. Schwamborn et al. (2011) provided students with a toolbar containing all relevant pictorial items in a computer-based experiment and instructed them to move and combine the elements by means of 'drag and drop' on a partly pre-drawn background for generating a picture of the learning content. After selecting an element, elements were replaced and could be used as often as wanted. Students who engaged in this kind of drawing construction scored better on a drawing test than students who were not instructed to generate drawings. It can be argued that this approach reduces potentially unnecessary reasoning regarding the visual appearance of the drawings' elements but at the same time, it constrains learners' creativity while drawing.

In another set of experiments, learner-generated drawing was investigated in the context of student's learning from oral prose (Lesgold et al., 1977, 1975). First-graders were given cutout figures of story characters and background scenes. While students listened to a prose story, they were instructed to illustrate what happened in the stories by organizing the cutout figures. This learner-generated drawing activity facilitated learning indicated by higher recall of story propositions only when students were provided with the correct pieces for the drawings but not when students had to select the pieces for each drawing out of a pool of cutouts. When children selected the pieces for their drawings out of a pool of items including distractors, drawing had either no or detrimental effects on learning. Although this task has been shown to be too difficult for young children (Lesgold et al., 1975), it can be argued that learners who are older should not experience this selection process as too demanding; rather they might experience increased cognitive activation that is helpful for learning. Thus, distractor items that were not useful for constructing correct drawings were included in Study 3 of the present thesis.

To present, there have been only few studies that have compared different forms of support varying with respect to the processes for which scaffolding is provided. To investigate the influence of different processes that are being scaffolded on learning, Study 3 compared learners who received baseline instructional support to learners who received extensive instructional support or no instructional support at all during drawing. In particular, baseline support consisted of providing a coarse layout of the drawing consisting of elements that were less relevant to understanding, so learners could focus on the appearance and spatial relations of relevant objects. During extensive support, learners were scaffolded by giving them all elements of the drawing, thereby constraining learners' reasoning about the visual-appearance these elements. Moreover, a control condition was included in which learners only read a text and did not engage in drawing at all.

### 1.5.4 Delayed testing

So far in the literature, effects of drawing have been investigated only for performance measures that were assessed immediately after learning. However, immediate assessments may not fully reflect the full potential of drawing as a learning strategy. Drawing comprises at least three cognitive processes – generation of new information, visualization, and externalization – that can all be considered challenging for learning, thereby possibly revealing relatively little benefit in tests administered immediately after learning. On the other hand, in line with the desirable difficulties approach (Bjork & Bjork, 2011; Bjork, 1994), challenging conditions may lead to more robust and flexible learning, which will be reflected in delayed rather immediate knowledge tests. When learners are faced with a difficulty that helps them to more deeply elaborate the learning content, it enhances long-term retention and the flexibility to transfer that knowledge to new contexts (Kerr & Booth, 1978; Rohrer & Taylor, 2007; Shea & Morgan, 1979). It can be argued that drawing might be considered a desirable difficulty because learners need to invest more resources during learning than simply perceiving the materials to construct the drawing, which consequently leads to deeper elaboration. When compared with learners who engage in a less cognitively demanding learning strategy, learners pursuing a learner-generated drawing strategy should benefit more from drawing in a delayed posttest, resulting in a larger gap between those strategies in a delayed posttest compared with an immediate posttest. This should be particularly true for tests measuring comprehension or transfer.

## 1.6 The Influence of Perceived Difficulty on Learner-Generated Drawing

According to research on drawing, it can be argued that a free-hand, unsupported drawing instruction might lead learners to experience difficulties in managing the mechanics of drawing itself, thereby reducing cognitive resources for meaningful learning which, in turn, decreases benefits of drawing (Schmeck et al., 2014; Schwamborn et al., 2010). To antagonize this risk, support like cutout figures or drawing prompts offers sufficient constraints and possibly leaves enough cognitive resources for learners to benefit from drawing through cognitive offloading. External representations are assumed to reduce the amount of cognitive effort which is required to solve a certain task (Scaife & Rogers, 1996). Nevertheless, generating external representations can have detrimental effects on text comprehension mediated by increased cognitive demands when no instructional support is provided (Leutner et al., 2009). It seems that offloading does not occur during free-hand drawing because of increased cognitive demands induced by the logistics of managing one's own drawing activity. On the other hand, research on drawing with instructional support showed that learners who were instructed to generate drawings and received instructional support perceived the learning material as equally difficult as learners who were instructed to engage in a text-focused strategy (Schmeck et al., 2014; Schwamborn et al., 2011). These findings indicate that mental offloading may have occurred during supported drawing as this promoted appropriate active processing while reducing extraneous cognitive processing in terms of perceived difficulty of the learning material (Sweller, 1999, 2005a). Hence, free-hand drawing requires a greater computational effort that should manifest itself in higher perceived difficulty, whereas instructional support should not only influence learning outcomes but also reduce perceived difficulty.

Therefore, the present thesis investigated the influence of perceived difficulty on learner-generated drawing. In particular, it was examined how drawing and other learning strategies affect perceived difficulty of the learning material and whether drawing aids can reduce perceived difficulty.

## 1.7 Overview of the Dissertation's Research Questions and Empirical Studies

In the present thesis, three major research questions were investigated addressing the role of contributing factors for drawing, the boundary conditions underlying benefits of drawing, and the influence of perceived difficulty on drawing. In the following, the three major research questions will be derived from previous research that was introduced earlier and a brief overview over the three empirical studies conducted within this dissertation is given. A more detailed derivation of hypotheses can be found in the respective chapters of the empirical studies within the following chapters of this thesis (see Chapter 2 for Studies 1 and 2, and Chapter 3 for Study 3).

The first research question focuses on what factors mainly contribute to benefits of drawing for learning outcomes. Based on previous research, one can argue that there may be at least three contributing factors to the benefits of drawing: (1) drawing involves active self-generation of a product that goes beyond what is explicitly stated in the learning material (cf. generation effect; Chi, 2009; Foos et al., 1994; Slamecka & Graf, 1978), (2) additional visualization of a verbally presented learning content (cf. dual coding theory; Clark & Paivio, 1991; Paivio, 1986), and (3) externalization of a previous mentally constructed (visual) representation (cf. multimedia effect; Butcher, 2014; Mayer, 2009). To disentangle these different contributing factors, and to get an insight on which factor may be most important for the benefits of a drawing activity, drawing was compared with other learning strategies that rely on only one or two of these factors. In Study 1, drawing was compared with strategies that either involved externalization and visualization but no generation (multimedia learning), externalization and generation but no visualization (writing summaries) or that involved none of these factors (learning with text only). In Study 2, drawing was contrasted with observing the pictorial representation of a multimedia presentation evolve step by step, which involves visualization and externalization but no generation. On the other hand, drawing was compared with mentally imagining the learning content, which involves visualization and generation but no externalization. If generation was the main contributing factor drawing should outperform strategies that differ from drawing with respect to generation. If visualization was the main contributing factor constructing drawings should lead to higher performance than strategies that do not involve dual coding of the learning material. If externalization was the main contributing factor a drawing strategy instruction should result in better learning than learning without an additional externalized representation. An interplay of all three factors should result in a benefit of drawing over all other learning strategies investigated in the present thesis.

The second research question is concerned with boundary conditions affecting the drawing strategy's benefit. In particular, it was investigated whether (1) benefits of drawing are more likely to be detected in a well-matched posttest, (2) whether the prognostic drawing effect is also influenced by the type of posttest, (3) whether test delay affects positive effects of drawing, (4) and whether different types of instructional support can have different effects on learning through drawing. First, based on previous research and on assumptions of the CMDC benefits of drawing should be observed in tests assessing higher-order knowledge but not in tests assessing lower-order knowledge (Leutner & Schmeck, 2014; van Meter & Firetto, 2013; van Meter & Garner, 2005). This assumption was tested in all three studies conducted within the present thesis. Second, the size of the prognostic drawing effect (Schwamborn et al., 2010) should be affected by the type of posttest in that the correlation between drawing quality and test performance should be higher for tests assessing higher-order knowledge than for tests assessing lower-order knowledge. Again, this assumption was examined in all three empirical studies. Third, it was of interest whether benefits of drawing are more pronounced in a delayed compared to an immediate posttest. In line with desirable difficulties research (Bjork & Bjork, 2011; Bjork, 1994), a larger gap between drawing and less demanding strategies should be observed in a delayed compared with an immediate posttest assessing higher-order knowledge. Long-term effects of drawing were investigated in Study 1 and Study 2. Fourth, it was of interest whether and how different forms of instructional support affect learning through drawing. To this end, two types of drawing aids (low vs. high support) were compared that differed with respect to the processes for which drawing was scaffolded. Low support was assumed to constrain reasoning about less relevant elements of the drawings, thereby enabling learners to fully concentrate on the visual appearance of and spatial relations between relevant elements. High support was assumed to free learners from reasoning about the visual appearance of any element of the drawings, thereby helping learner to fully focus on the spatial relations between those elements.

The third research question deals with whether a drawing activity positively or negatively influences perceived difficulty of the learning material. Drawing may reduce the perceived difficulty of the learning material because mental offloading occurs during learning with external representations (Scaife & Rogers, 1996). If this assumption holds true, there should be no differences between drawing and other strategies that involve learning with external representations. However, constructing drawings can have detrimental effects on text comprehension mediated by increased cognitive demands when no instructional support is provided (Leutner et al., 2009). Offloading does not seem to occur during free-hand drawing because of increased cognitive demands induced by the logistics of managing one's own drawing activity. But when instructional support is provided during drawing, mental offloading may occur as this promotes appropriate active processing while reducing extraneous cognitive processing in terms of perceived difficulty of the learning material (Schmeck et al., 2014; Schwamborn et al., 2011; Sweller, 1999, 2005a). Hence, free-hand drawing was assumed to require a greater computational effort that should be reflected in higher perceived difficulty, whereas instructional support should not only influence learning outcomes positively but also reduce perceived difficulty. This research question was addressed in the Studies 2 and 3.

Study 1 investigated which factors mainly contribute to benefits of drawing while focusing on the comparison of generation with visualization, and whether drawing also affects performance in a delayed posttest. A condition in which students were asked to generate drawings for each paragraph of an expository text during learning was compared with a multimedia condition in which students learned with a combination of text and pictures, a summary condition in which students were asked to generate a written summary of the main concepts, and a text-only condition in which students were provided only with a text from which to learn. After learning, learning outcomes were assessed with an immediate and a delayed posttest. Time of testing was manipulated within subjects.

Study 2 directly builds up on Study 1 also investigating what factors mainly contribute to the benefits of drawing and possible long-term effects. Thereby, Study 2 focused on the factors generation and externalization. Drawing was compared with an observation condition in which students were provided with a text and dynamic visualizations showing a step-by-step evolution of a drawing, and a mental imagery condition in which students were asked to mentally imagine the content of each paragraph. After learning, learning outcomes were assessed with an immediate and a delayed posttest. Time of testing was manipulated between subjects.

Study 3 examined the influence of different types of instructional support on the effectiveness of drawing. Therefore, a no-support condition in which learners received a drawing instruction but no additional support was contrasted to a low-support condition, a high-support condition, and a text-only control condition. In the low-support condition, learners received drawing instructions and pre-drawn backgrounds for drawing construction, so that the requirement to reason about the appearance and spatial relations of irrelevant objects was reduced. In the high-support condition, learners received drawing instructions and a box containing different elements and labels to use for drawing construction that should reduce unnecessary reasoning about the visual appearance of the drawings' elements, thereby enabling learners to fully focus on the spatial relations between relevant elements during drawing. Learning outcome performance was assessed immediately after learning.

# 2 Which Factors Do Mainly Contribute To Effects of Learner-Generated Drawing? Investigating the Influence of Visualization, Generation, and Externalization

The present chapter covers a study containing two experiments (Study 1 and Study 2). The present study investigated what factors mainly contribute to benefits of a drawing strategy instruction. Also, boundary conditions that influence the effectiveness of drawing were of interest. In particular, I aimed at replicating previous findings that benefits of drawing depend on the type of posttest used to assess learning outcome performance in that benefits of drawing should be revealed in assessments of higher-order knowledge but not in assessments of lower-order knowledge. Moreover, the dependence of the prognostic drawing effect on the type of posttest as well as the influence of test delay on learning was examined. Lastly, it was investigated how free-hand drawing relates to perceived difficulty. Thus, all three research questions of the present thesis were addressed. Study 2 directly builds up on Study 1, so both experiments are reported and discussed within one chapter.

## 2.1 Overview and hypotheses

Drawing constitutes a learning strategy during which learners rely on a written text to construct representational visualizations that depict the key elements and their relations described in that text (Alesandrini, 1984; Schmeck et al., 2014; van Meter, 2001). During drawing, learners engage in generative learning processes while constructing a representation of the learning content that goes beyond what is explicitly stated in the written text, which may result in a deeper understanding of the learning content (Wittrock, 1990). Through drawing, learners are furthermore assumed to create an internal dual code of the learning content, where the written text yields a linguistic mental representation while the learner-generated drawing represents a pictorial code (Paivio, 1986). Finally, learner-generated drawing results in a multimedia representation (Mayer, 2014a) of the learning content because the learner's mental representation of the text is externalized onto paper. In sum, there are at least three factors that may play a role during drawing: generation, visualization, and externalization.

In the present study, drawing was compared with other learning strategies that differ from drawing with respect to these factors. This was done to investigate the question of what mainly contributes to the benefit of drawing. Moreover, it was examined whether these factors differ in terms of sustainable knowledge that is assessed after a delay rather than immediately after learning.

Two experiments were conducted to investigate effects of drawing on immediate and delayed learning outcome performance after learning from an expository text about biomechanics in human swimming behavior. There is a lack of research on which of the factors of generation, visualization, and externalization mainly drive benefits of learner-generated drawing. Accordingly, the two experiments aim to investigate this issue in a systematic way. To this end, in Study 1, it was varied whether learners applied a strategy that involves visualization and/or engaged in a generative learning activity. Specifically, a drawing condition was contrasted with a multimedia, a summary-writing, and a text-only condition. Because Study 1 revealed visualization (independent from generation) to be a driving factor for the beneficial effects of drawing, Study 2 focused on strategies relying on this factor. In Study 2, a closer look was taken at the role of externalization within strategies that all used visualization; more precisely, it was investigated whether self-generating a visualization or externalizing a generated visualization is crucial for beneficial effects of drawing. Hence, in Study 2, a drawing condition was contrasted with an observation of a step-by-step multimedia presentation condition and a mental imagery condition.

Four research questions guided the present study. First, it was of interest what factors mainly contribute to benefits of drawing for learning outcomes. In line with previous research and predictions by the CMDC (van Meter & Firetto, 2013), no differences between drawing and other learning strategies for lower-order knowledge assessments of recognition performance were expect, as was for assessments of higher-order knowledge (Hypothesis 1). Hence, the following hypotheses focus on assessments of transfer and visuo-spatial knowledge. If generation is the main contributing factor, the drawing condition should outperform the multimedia, text-only, and observation conditions; moreover, there should be no differences between the drawing, summary, and mental imagery conditions (Hypothesis 1a). If visualization is the main contributing factor, the drawing condition should outperform the summary and text-only conditions and should perform equally well as the multimedia, observation, and mental imagery conditions (Hypothesis 1b). If externalization is the main contributing factor, learners in the drawing condition should outperform learners in the text-only and mental imagery conditions, while there should be no differences between the drawing, multimedia, summary, and observation conditions (Hypothesis 1c). If all three investigated factors contribute to the drawing effect, then the drawing condition should outperform all other conditions (Hypothesis 1d).

Second, it was of interest whether benefits of drawing are more pronounced in a delayed compared to an immediate posttest. In line with desirable difficulties research, a larger gap was expected between drawing and less demanding strategies in a delayed compared with an immediate posttest for the higher-order assessments (transfer and visuo-spatial knowledge; Hypothesis 2a). For the lower-order assessments (recognition performance) forgetting over time was expected for all learning strategies reflected by higher test performance in the immediate than in the delayed posttest (Hypothesis 2b).

Third, the relation between the quality of the constructed drawings in the drawing condition and learning outcome performance was investigated. Based on previous research on the prognostic drawing effect (Schmeck et al., 2014; Schwamborn et al., 2010), the correlation between drawing quality and learning outcome performance was expected to be positive (Hypothesis 3a). Furthermore, in line with the CMDC's assumptions (van Meter & Firetto, 2013) different relations between drawing quality and learning outcome performance were expected depending on the type of posttest. More precisely, larger positive correlations between drawing quality and learning outcome performance were expected for posttests assessing higher-order knowledge like transfer and visuo-spatial knowledge than for posttests assessing lower-order knowledge like recognition (Hypothesis 3b).

Fourth, it was examined how perceived difficulty of the learning material was affected by drawing. This research question was tested only in the second experiment. It can be argued that constructing drawings reduces the perceived difficulty of the learning material because mental offloading occurs during learning with combinations of text and pictures (Sweller, 1999, 2005a). Thus, if this assumption holds true, there should be no differences between drawing, mental imagery, and observation in subjective measures of perceived difficulty (Hypothesis 4a). On the other hand, one can argue that a drawing instruction increases perceived difficulty of the learning material by means of increased cognitive demands induced by the logistics of managing the drawing process (Leutner et al., 2009). Thus, it would be expected that the drawing condition reports higher values of perceived difficulty than the imagery and observation conditions (Hypothesis 4b).

## 2.2 Study 1

### 2.2.1 Method

**Participants and design.** Participants were 121 undergraduate students from a university in the southwest of Germany (88 females; *M* = 22.71 years, *SD* = 3.60) who took part in the experiment voluntarily for an incentive of 16 euros or course credit. Biology, medicine, physics, and sports sciences majors were excluded from the experiment to avoid high prior knowledge of the learning content (biomechanics in human swimming behavior). Participants were randomly assigned to one of four experimental conditions that resulted from a 2x2x2 mixed design with the between-subjects factors ‘visualization’ (with vs. without) and ‘generation’ (with vs. without) and the within-subjects factor ‘test time’ (immediate vs. one week delay). The four experimental conditions were drawing (with visualization, with generation; $n$ = 31), multimedia (with visualization, without generation; $n$ = 30), summary (without visualization, with generation; $n$ = 30) and text-only (without visualization, without generation; $n$ = 30). The latter served as a control condition. All conditions with the visualization factor level ‘with’ also included externalization.

**Materials.** A text about biomechanics in human swimming behavior was used as the learning material, which contained 976 words in 9 paragraphs. Each paragraph was presented on a separate page. The paragraphs were about muscle contraction (111 words), series connection of muscle fibers (100 words), parallel connection of muscle fibers (90 words), pennate muscles (110 words), buoyancy (126 words), the human body in water (116 words), fluid resistance (103 words), the drive concept of action-reaction (89 words), and the drive concept of hydrodynamic lift (131 words). The layout of the pages depended on experimental condition (see Figure 4). In the text-only condition, only text was presented in the middle of each page aligned to the top. In the other three conditions, the text was presented on the right of each page aligned to the top. In the drawing condition, the left side of each page was blank with a prompt asking participants to construct their drawing there. In the multimedia condition, additional pictures were presented on the left aligned to the center, illustrating the structures and their spatial relations described in the text. In the summary condition, the left side of each page was blank with a prompt for participants to write their summary there.

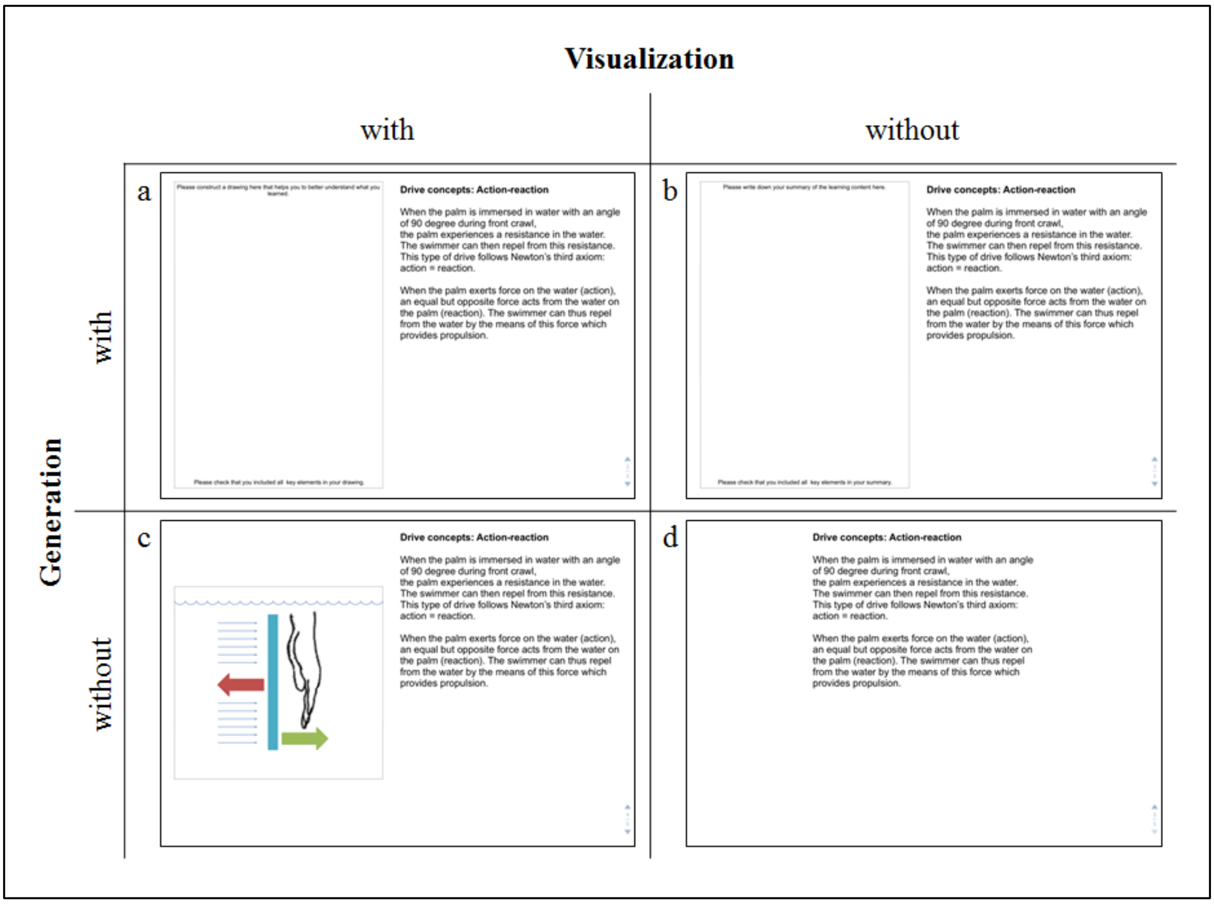


*Figure 4.* Sample page of the learning material used in Study 1 (translated from the German original) depending on the experimental condition: (a) drawing, (b) summary, (c) multimedia, (d) text-only.

**Measures.** Verbal ability, spatial ability, prior knowledge, and interest in the learning topic were assessed as control variables. The dependent variable was learning outcome as assessed in the immediate and delayed posttest. Additionally, drawing quality was assessed for the drawings constructed during learning in the drawing condition.

Verbal ability was assessed using the reading comprehension measure of the standardized instrument LGVT 6-12 (Schneider, Schlagmüller, & Ennemoser, 2007). This test consists of a text with 23 multiple-choice cloze items. Participants chose from 3 alternative words the one that fit the meaning of the cloze sentence (retest reliability $r$ = .87). To assess spatial ability, 10 multiple-choice paper-folding items were used taken from the Paper Folding Test (PFT; Ekstrom, French, & Harman, 1976; Cronbach's alpha = .64). 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 myofibrils.”; Cronbach’s alpha = .04). Each item consisted of a short statement that was a non-verbatim version of the learning text with a two-choice answer format (right vs. wrong). Interest was measured using the corresponding subscale of a short scale of intrinsic motivation (KIM; Wilde, Bätz, Kovaleva, & Urhahne, 2009) that consisted of three items. An example item (translated from the German original) is “I enjoyed how I was working with the learning content.” Each item was rated on a 5-point Likert-scale ranging from *full agreement* to *no agreement at all* (Cronbach’s alpha = .89).

Learning outcome was assessed with a recognition test, transfer test, and a drawing test. The recognition test consisted of the same 20 verification items used in the prior knowledge test and assessed participants’ recognition of the learning content (Cronbach’s alpha = .46). For computing the recognition test scores (prior knowledge and posttest) each participant was awarded with 1 point for each correct answer; the sum of correctly answered items was transformed into a percentage for easier interpretation. Because one of two answer choices was correct in every verification item, guessing the answers would yield a score of 50% by chance only. Thus, recognition and prior knowledge test scores were adjusted for guessing probability so that 0 % means ‘no knowledge’ and 100 % means ‘perfect knowledge’. This was done by subtracting 50 from each participant’s score and multiplying the resulting value by 2 ([x – 50] *x* 2). The transfer test consisted of seven open-ended questions used to assess the participants’ ability to apply their knowledge to novel contexts (e.g., “How can you explain the lift of an aircraft?”; Cronbach’s alpha = .40). The transfer test scores were computed by awarding each participant with 1 point for each correct answer and transforming into a percentage. The drawing test consisted of eight drawing items. Participants answered open-ended questions by drawing, by correcting/completing supplied incorrect or partial drawings, and by labeling parts of supplied drawings (e.g., “Please draw in the missing fibers of a muscle with series connection."; Cronbach's alpha = .58). This test assessed participants' visuo-spatial knowledge of the learning content, whereby the possible number of points differed between the items according to the number of main ideas 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 21 % of the data for the transfer and drawing tests. Inter-rater agreement was high for both the transfer test (all Krippendorff's alphas > .94), and the drawing test (all Krippendorff's alphas > .91) so that one of the raters scored the remaining data.

The quality of the drawings that participants in the drawing condition constructed for each paragraph during learning was also assessed. To this end, the main concept units within each paragraph were identified and it was assessed whether these were correctly incorporated into the drawings or not present at all. For the drawing accuracy score, participants received one point for each correctly drawn main concept unit, yielding a maximum score of 76 points. Then, the percentage correct was calculated. For the drawing omissions score, one point was given for each omitted main concept unit, yielding again a maximum score of 76 points. Similar to the drawing accuracy score, the percentage was calculated. Drawing quality was assessed only for participants in the drawing conditions. Two raters scored 23 % of the data for drawing quality. Inter-rater agreement was acceptable for both drawing accuracy (all Krippendorff's alphas > .86), and drawing omissions (all Krippendorff's alphas > .89) so that one of the raters scored the remaining data.

**Apparatus.** Each participant was provided with a tablet computer (Apple iPad 4 with an integrated 9.7'' touchscreen) on which all materials and measurements were presented, except for the LGVT 6-12, PFT, and the instructions for the learning episode, which were presented in a paper-based format. All measures, except for the drawing test, were web-based. The presentation of the learning material in all four conditions and the drawing test was implemented using the application "Notability" (Ginger Labs Inc., version 5.33), which allows both typing and the construction of drawings. A toolbar at the top of the screen contained several functions that participants were able to use while working on the learning material including 'undo' (undo the last step), 'write text' (entering text via the keyboard of the tablet-computer; font type, style and size could be adjusted), 'crayon' (drawing; color and thickness of lines could be adjusted), 'highlighting' (drawing transparent, thicker lines; color and thickness of lines could be adjusted), 'eraser' (erasing what was drawn before with a touch on the relevant position) 'cut and paste' (selecting parts of 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.

**Procedure.** Participants were randomly assigned to one of the four experimental conditions and were invited to two experimental sessions separated by one week. Participants were tested in groups with a maximum of six participants per session. Each participant was seated in a semiprivate cubicle in front of a tablet computer. First, reading comprehension and spatial ability were assessed. Then, participants were instructed to switch to the tablets and to work individually on some demographic questions and the prior knowledge test. This was followed by instructions for the learning episode, which included a list of the available functions of 'Notability' and explanations on how to use them as well as detailed instructions on how to work on the learning material according to the experimental condition they were assigned to. All participants were first told that the strategy they should engage in has proven to be effective for learning. They then received a list of steps they should follow to successfully carry out the respective learning strategy. Participants in the drawing condition were instructed to thoroughly read the text and to construct a drawing for each paragraph that contained all the important information of the text, whereby the drawings did not need to look pretty. Participants in the multimedia condition were instructed to thoroughly process the text and the pictures and to identify the most important information of the text in the pictures. Participants in the summary condition were told to thoroughly read the text and to construct for each paragraph a written summary that contained all the important information of the text. Participants in the text-only condition were instructed to thoroughly read the text. Before participants were provided with the learning material, they worked on a short task to familiarize themselves with “Notability”. Participants then could navigate through the learning material in a self-paced manner with the ability to move forwards and backwards. The learning episode was followed by the measurement of interest. After that, the immediate posttest was administered. Participants then were dismissed and instructed to come back to the lab one week later to work on the delayed posttest. During the second session, participants completed the delayed posttest and then were debriefed, paid and dismissed.

### 2.2.2 Results

**Control variables.** Means and standard deviations are reported in Table 2. Before testing the hypotheses, it was analyzed whether the four experimental groups differed in gender, age, verbal ability, spatial ability, prior knowledge, or interest. A chi-square analysis revealed no significant differences regarding gender ($p$ = .504). To analyze the remaining five control variables a multivariate analysis of variance (MANOVA) was conducted with the between-subjects factors visualization (with versus without) and generation (with versus without). It revealed no significant main effect of visualization, Pillai’s trace = 0.075, $F$(5,112) = 1.81, $p$ = .12, partial eta$^2$ = .08, and the main effect of generation as well as the visualization x generation interaction were not significant, both $F$s < 1. Hence, conditions were similar with regard to these entry characteristics.

Table 2

*Means and standard deviations for the control variables as a function of experimental condition in Study 1*

| | Visualization | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | With | | | | Without | | | |
| | With Generation | | | | Without Generation | | | |
| | With (Drawing) | | Without (Multimedia) | | With (Summary) | | Without (Text-only) | |
| *n* | 31 | | 30 | | 30 | | 30 | |
| | *M* | *SD* | *M* | *SD* | *M* | *SD* | *M* | *SD* |
| Age | 22.77 | 3.32 | 22.17 | 3.40 | 22.53 | 3.69 | 23.37 | 4.01 |
| Verbal ability (0-100) | 61.03 | 10.16 | 61.63 | 11.74 | 63.70 | 9.69 | 62.67 | 12.22 |
| Spatial ability (-10-10) | 5.65 | 2.84 | 6.00 | 2.82 | 5.37 | 3.14 | 5.60 | 3.21 |
| Prior knowledge (0-100%) | 5.48 | 27.43 | 10.33 | 18.29 | 9.00 | 20.23 | 3.33 | 15.61 |
| Interest (1-5) | 4.02 | 0.78 | 3.97 | 0.71 | 3.67 | 0.87 | 3.53 | 0.96 |

Table 3

*Means and standard deviations for the learning outcome measures as a function of experimental condition and time of testing (immediate vs. delayed) in Study 1*

| Time of testing | Visualization | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | With | | | | Without | | | |
| | With Generation | | | | Without Generation | | | |
| | With (Drawing) | | Without (Multimedia) | | With (Summary) | | Without (Text-only) | |
| *n* | 31 | | 30 | | 30 | | 30 | |
| | *M* | *SD* | *M* | *SD* | *M* | *SD* | *M* | *SD* |
| Immediate | | | | | | | | |
| Recognition (% correct) | 17.58 | 15.59 | 15.33 | 14.91 | 12.83 | 11.94 | 21.40 | 12.23 |
| Transfer (% correct) | 55.30 | 20.45 | 54.76 | 22.22 | 42.86 | 23.13 | 56.09 | 25.26 |
| Drawing (% correct) | 46.54 | 15.46 | 66.52 | 22.08 | 37.65 | 22.36 | 47.64 | 15.38 |
| Delayed | | | | | | | | |
| Recognition (% correct) | 12.71 | 13.43 | 11.59 | 15.49 | 9.50 | 14.28 | 16.17 | 14.24 |
| Transfer (% correct) | 60.05 | 13.47 | 59.09 | 25.09 | 26.19 | 22.41 | 52.86 | 21.92 |
| Drawing (% correct | 50.49 | 17.45 | 65.87 | 21.43 | 40.36 | 19.83 | 46.75 | 15.60 |

**Learning outcomes.** Means and standard deviations are reported in Table 3. Before analyzing the recognition performance difference scores were computed by subtracting the scores of the prior knowledge test from the immediate recognition posttest scores and from the delayed recognition posttest scores. When positive, these difference scores reflect the amount of knowledge gain between the prior knowledge test and the immediate or delayed posttest. This approach was used to test knowledge gains between prior knowledge test and posttest scores and differences between conditions within one test. Then, a repeated measures analysis of variance was conducted with the immediate and delayed difference scores as the within-subjects factor and visualization and generation as the between-subjects factors. It revealed that the difference scores deviated significantly from zero, $F(1,117) = 155.26$, $MSE = 334.00$ $p < .001$, partial eta² = .57, indicating that participants reached significant knowledge gains between the prior knowledge test and the immediate and the delayed posttest. Moreover, it revealed a significant main effect of time, $F(1,117) = 22.74$, $MSE = 49.08$, $p < .001$, partial eta² = .16, showing that participants performed less well in the delayed posttest ($M = 12.49$ %, $SD = 13.94$) than in the immediate posttest ($M = 16.79$ %, $SD = 13.97$). The visualization x generation interaction was also found to be significant, $F(1,117) = 3.92$, $MSE = 334.00$, $p = .050$, partial eta² = .03 (see Figure 5). Post-hoc multiple comparisons revealed that participants in the without generation condition (text-only, $M = 18.76$ %, $SE = 2.36$) scored higher than participants in the generation condition (summary, $M = 11.17$ %, $SE = 2.36$; $p = .02$) when visualizations were absent, but that this difference could not be found when visualizations were present ($p = .61$). The main effect of generation did not reach significance, $F(1,117) = 1.59$, $MSE = 334.00$, $p = .21$, partial eta² = .01. The main effect of visualization and all other interactions were not significant, $Fs < 1$.

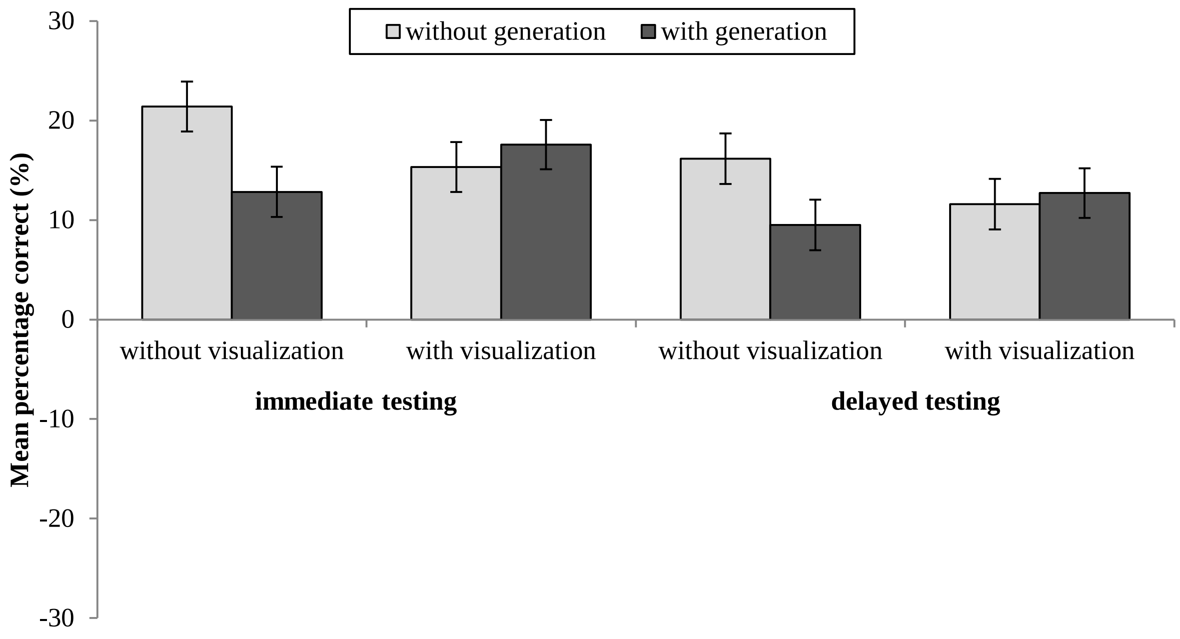


*Figure 5*. Mean differences (% correct) representing recognition test performance subtracted from the prior knowledge test as a function of experimental condition. Error bars represent standard errors.

For transfer performance, a repeated-measures ANOVA was conducted with the immediate and delayed transfer scores as the within-subjects factor and visualization and generation as between-subjects factors. It revealed a significant main effect of time showing that overall, participants reached higher scores in the delayed (*M* = 54.59 %; *SD* = 21.59) than in the immediate posttest (*M* = 49.80 %; *SD* = 23.17), $F(1,117) = 8.23$, $MSE = 168.93$, $p < .01$, partial eta² = .07. Moreover, a significant main effect of visualization was found, $F(1,117) = 8.05$, $MSE = 797.70$, $p < .01$; partial eta² = .06. The groups with visualization (drawing and multimedia; *M* = 57.31 %; *SD* = 18.92) scored higher on the transfer items than those without visualization (summary and text-only; *M* = 47.00 %; *SD* = 20.81). The main effect of generation and all interactions were not significant, all $Fs < 1$.

For the performance on the drawing test, a similar repeated measures ANOVA was conducted. A similar pattern as for the transfer test was found, namely, that the groups with visualization (*M* = 57.21 %; *SD* = 19.83) scored higher than the groups without visualization

(*M* = 43.10 %; *SD* = 17.83), $F(1,117) = 19.66$, *MSE* = 625.48, $p < .001$, partial eta² = .14. Additionally, a significant main effect of generation revealed that the without generation groups (multimedia and text-only; *M* = 56.69 %; *SD* = 19.83) scored higher than the generation groups (drawing and summary; *M* = 43.84 %; *SD* = 18.35), $F(1,117) = 16.17$, *MSE* = 625.48, $p < .001$, partial eta² = .12. The interaction between the factors time of testing and generation was at the verge of significance, $F(1,117) = 2.86$, *MSE* = 88.97, $p = .09$, partial eta² = .02. The main effect of time of testing was not significant, $F(1,117) = 1.11$, *MSE* = 88.97, $p = .29$, partial eta² = .01. The visualization x generation interaction was also not significant, $F(1,117) = 2.18$, *MSE* = 625.48, $p = .14$, partial eta² = .02, nor were the remaining interactions, all *F*s < 1.

**Drawing quality.** To test Hypotheses 3a and 3b, the quality of the learner-generated drawings constructed in the drawing condition was analyzed. An exploratory analysis showed that drawing accuracy scores ranged from 16.86% to 59.39% (*M* = 37.86%, *SD* = 9.67). Means for drawing accuracy scores were significantly different from both the bottom (0%; $t(30) = 21.80$, $p < .001$) and the top of the scale (100%; $t(30) = 35.79$, $p < .001$). Drawing omissions scores ranged from 35.01 to 74.31% (*M* = 51.77%, *SD* = 9.36). Means for drawing omissions scores were significantly different from both the bottom (0%; $t(30) = 30.79$, $p < .001$) and the top of the scale (100%; $t(30) = 28.68$, $p < .001$). Thus, participants produced drawings of medium quality. There were no floor or ceiling effects. Different from what would have been expected according to the prognostic drawing effect, neither drawing accuracy nor drawing omissions scores were significantly correlated with any of the learning outcome measures in the immediate or delayed posttest (all *r*s < .20, all *p*s > .30).

### 2.2.3 Summary and conclusions

The beneficial effects of drawing and multimedia material presentation on transfer and visuo-spatial knowledge support the predictions of Hypothesis 1b that learning with a linguistic and an additional pictorial code aids learning (Paivio, 1991). These results indicate that instructing learners to generate representational drawings supported them in building a coherent mental model of the relevant objects and their spatial relations equally well as does providing them with multimedia material. It is yet an open question, however, whether this visualization needs to be available as an external representation, which was addressed in Study 2. The finding that the drawing and the multimedia groups outperformed the summary and the text-only groups on the transfer and drawing tests but not on the recognition test is in line with the prediction of Hypothesis 1 and with research on learner-generated drawing showing that benefits of drawing are more pronounced on measures that reflect higher-order knowledge (Leopold & Leutner, 2012; van Meter & Firetto, 2013).

Unexpectedly, no benefits of generative activity were found. The results indicate that engaging in a generative learning activity did not enhance learning and actually hindered learning with respect to visuo-spatial knowledge, contrary to the generation effect (Chi & Wylie, 2014; Foos et al., 1994; Slamecka & Graf, 1978). In line with previous studies, this finding suggests that when engaging in a generative activity, learners' attention might be focused mainly on the construction of a summary or drawing and thus, fewer resources are left available to engage in the construction of a mental model that contains structural objects and their spatial relations (Leopold & Leutner, 2012). This lack of mental model construction might be more pronounced in a test that is sensitive to these spatial relations, like the drawing test used in the present experiment.

Due to the opposing findings regarding the generation effect, the second experiment focused mainly on the generation and externalization aspects of constructing drawings to gain more insight on which generation processes are most important for learner-generated drawing. To this end, three different variants of instruction were used in Study 2: drawing, mental imagery, and observation. Instead of comparing drawing with a multimedia condition using static pictures, an observation condition was included that included dynamic visualizations. In this condition, similar to the multimedia condition of Study 1, written text was accompanied by visualizations of the learning content, but these visualizations, rather than being static, evolved step by step, similar to how a drawing is constructed during learning. Thus, in this condition, participants observed the generation of a drawing instead of self-generating it. We also included a measure of perceived difficulty of the learning material in Study 2 to better understand the cognitive resources available for meaningful learning.

The finding that delayed test scores were higher than or equal to immediate test scores in all four groups for higher-order knowledge suggests that learners may have benefited from retrieval practice because we used the same set of items for the immediate and the delayed posttests (Roediger, Putnam, & Smith, 2011). It might hence be that the results were confounded with respect to the within-subjects manipulation of time of testing. To account for that, time of testing was included as a between factor in the second experiment. At present, there is no indication that generating an external representation is a desirable difficulty (Bjork & Bjork, 2011), since generation had no effect on differences between the immediate and delayed test. Hence, the results did not support the predictions of Hypothesis 2a.

In contrast to the predictions (Hypotheses 3a and 3b), previous findings in favor of a prognostic drawing effect could not have been replicated (Schmeck et al., 2014; van Meter & Firetto, 2013). That is, no relation was found between drawing quality of drawings constructed during learning and learning outcomes. However, participants in the drawing condition were able to benefit from this strategy to some extent as indicated by increased higher-order knowledge. The lack of a prognostic drawing effect in Study 1 might be explained by the spatial information in the learning text not being salient enough for participants to use it for constructing their drawings. This assumption is supported by the relatively low drawing accuracy (*M* = 37.86 %) and relatively high drawing omissions (M = 51.77 %) that was found in this experiment. To account for this, the learning text was edited and spatial information was made more salient in Study 2.

## 2.3 Study 2

### 2.3.1 Method

**Participants and design.** Participants were 223 undergraduate students from a university in the southwest of Germany, who took part in the experiment voluntarily for an incentive of 12 euros or course credit. Biology, medicine, physics and sports sciences majors were excluded from the experiment to avoid high prior knowledge of the learning content (biomechanics in human swimming behavior). In total 19 participants were excluded from the analyses: Eight participants did not follow the instructions during the learning episode correctly (e.g., by writing summaries instead of constructing drawings), six participants did not attend the second test session, two participants already knew the learning material used in this experiment, one participant had insufficient knowledge of German, one participant was provided with the wrong learning material, and one participant was a medicine major. Thus, the analyses were based on 204 participants (151 female; *M* = 23.15 years, *SD* = 2.88) Participants were randomly assigned to one of six experimental groups that resulted from a 3x2 between-subjects design with the factors ‘learning strategy’ (drawing vs. observation vs. imagery) and ‘time of testing’ (immediate vs. 1 week delayed). The six experimental groups were drawing immediate (*n* = 36), observation immediate (*n* = 34), imagery immediate (*n* = 34), drawing delayed (*n* = 37), observation delayed (*n* = 31) and imagery delayed (*n* = 32).

**Materials.** For the learning episode, a revised version of the text from Study 1 was used. Spatial information was explained in more detail and thus made more salient for learners, resulting in a slightly longer text compared to Study 1 (983 words in 9 paragraphs). In all three experimental conditions, the text was presented on the left of each page aligned to the top, while the right side of each page was blank. There, a prompt for participants was presented instructing them to either construct a drawing (drawing groups) or to mentally imagine the content of the paragraph (imagery groups). In the observation groups, author-generated drawings evolving step-by-step were presented as videos on the right side of each page. Each page of the learning material started with the text presented on the left. After a short time – an average reading time for the respective paragraph – the author-generated drawings started to build up on the right. The completed drawing was presented for several seconds before the next page of the learning material was presented. Participants were able to pause, fast-forward, and rewind the video (see Figure 6 for an example of different stages of an author-generated drawing for one paragraph in the observation conditions).

**Measures.** Verbal and spatial ability, prior knowledge and interest in the learning topic were assessed as control variables. Moreover, a mental imagery rating was included that served as a manipulation check regarding the mental imagery conditions. Learning outcome, as assessed in the immediate and delayed posttest, perceived difficulty of the learning material, and drawing quality of constructed drawings during the learning episode were assessed as dependent variables.

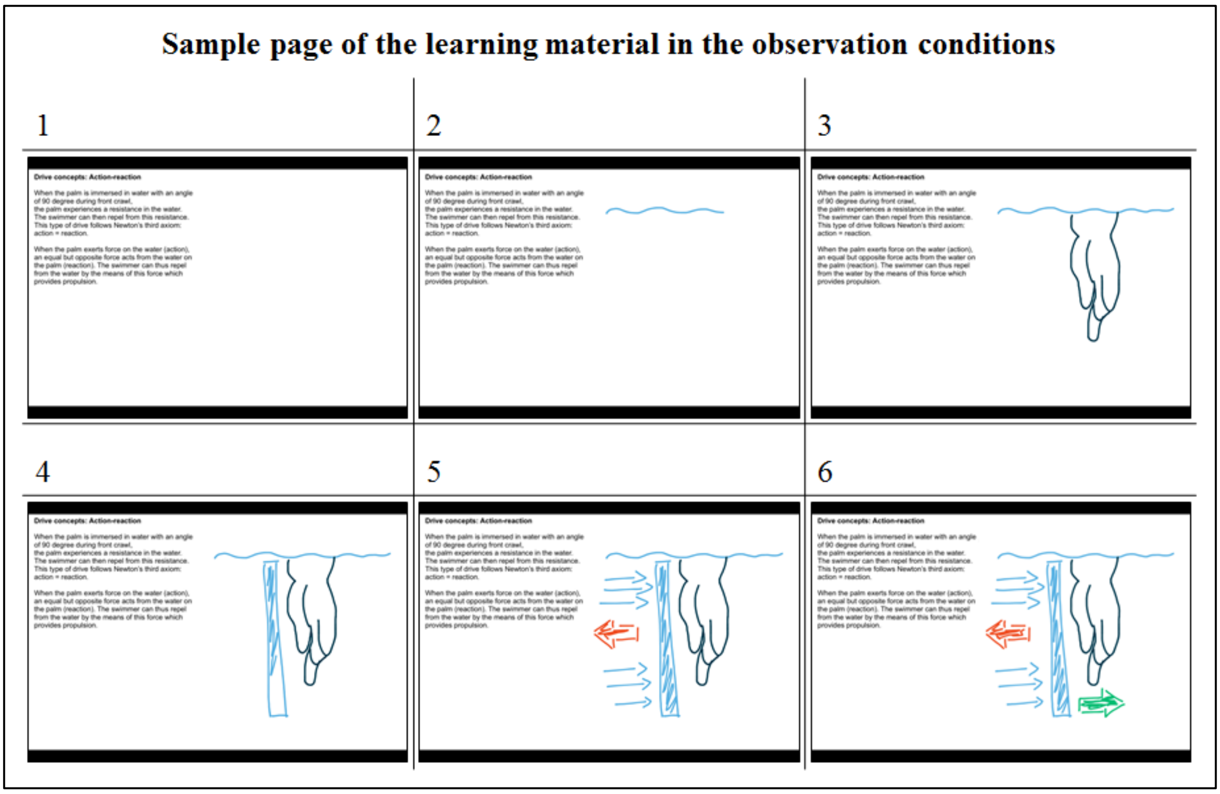


*Figure 6*. Sample page of the learning material (translated from the German original) used in the observation conditions of Study 2. The different stages of the author-generated drawings are shown, beginning with text only (1), evolving step-by-step (2-5) and resulting in the finished drawing (6).

Verbal ability was assessed using a computer-based adaption of the sentence finishing subtest of the I-S-T 2000 R (Liepmann, Beauducel, Brocke, & Amthauer, 2007) containing 20 multiple-choice items (Cronbach's alpha = .69). In this test, participants are instructed to choose one out of four words that fits the meaning of a cloze sentence. Spatial ability was assessed using a different 10-item subset of the PFT as in Study 1 in a computer-based version (Cronbach's alpha = .71). Prior knowledge was measured with the same items that were used in Study 1 (Cronbach's alpha = .04). Identically to Study 1, interest was assessed using the three KIM items (Cronbach's alpha = .82). The mental imagery rating consisted of five items (e.g., "I created mental images of the learning content.") that were rated on a 5-point Likert scale ranging from *not correct at all* to *fully correct* (Cronbach's alpha = .92) that were designed to capture how strongly the participants in the imagery conditions followed the instructions. Five additional items (e.g., “I tried to memorize key terms.”) served as filler items to disguise the purpose of the questionnaire and were not included in the analyses.

Learning outcome was assessed similarly to Study 1. The recognition test consisted of the same 20 verification items that were used in Study 1 (Cronbach’s alpha = .37). For the transfer test, the seven items of Study 1 and one additional item were used to represent each main concept of the learning text by the same number of items. Thus, the transfer test consisted of eight items in total (Cronbach’s alpha = .48), for which a maximum of 9 points could be earned. Ten items were used for the drawing test: eight items that were used in Study 1 and two additional items to capture each main concept of the learning text (Cronbach’s alpha = .68). For the drawing test, participants could earn a maximum of 36 points. All scores were transformed into percentages for easier interpretation. Two raters scored 24 % of the data for the transfer and drawing tests. Inter-rater agreement was high for both the transfer test (all Krippendorff’s alphas > .93), and the drawing test (all Krippendorff’s alphas > .90) so that one of the raters scored the remaining data.

Perceived difficulty was measured using three items (e.g., "How easy or difficult was it for you to learn about biomechanics in human swimming behavior?"; Paas, 1992) that were rated on a 7-point Likert scale ranging from *very easy* to *very difficult* (Cronbach’s alpha = .78).

Drawing quality of constructed drawings during learning was assessed as in Study 1, only for participants in the drawing conditions. Like in Study 1, two separate percentage scores were computed for drawing accuracy and drawing omissions. Two raters scored 22 % of the data for drawing quality. Inter-rater agreement was acceptable for both drawing accuracy (all Krippendorff’s alphas > .88), and drawing omissions (all Krippendorff’s alphas > .90) so that one of the raters scored the remaining data.

**Apparatus.** In all six experimental conditions, all materials were presented on a tablet computer (iPad 4.0 with iOS 9.0). Except for the drawing test, all measures were web-based. The learning material in the drawing and imagery conditions and the drawing test were presented using the application 'Notability' (Ginger Labs Inc., version 5.6.2). The learning material in the observation conditions was presented using the application 'Educreations' (Educreations Inc., version 2.0.11). With this application, the learning material was presented in a video format and participants were able to pause, fast forward, and rewind the learning material video. Each iPad was accompanied by a stylus pen to be used during the experiment.

**Procedure.** Participants were invited for two experimental sessions separated by one week. Upon arrival at the lab, they were seated in front of a tablet computer in a semiprivate cubicle. A maximum of six participants could be tested at the same time. Participants were randomly assigned to one of the six experimental conditions.

The procedure was similar to that of Study 1, except that the measurement of verbal and spatial ability was carried out after the learning episode. Participants started with some demographic questions and then received information about how to work on the learning material. The instructions followed the same structure as in Study 1, that is, participants were told that the strategy they should engage in has proven to be effective for learning. Moreover, participants in the drawing conditions were instructed to thoroughly read the text 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 observation conditions were instructed to thoroughly process the text and the evolving pictures and to identify the most important information of the text in the pictures. Participants in the imagery conditions were instructed to thoroughly process the text and then to mentally imagine the most important aspects of each paragraph. Afterwards, participants were provided with the learning material. After the learning episode, participants worked on the questionnaires to assess perceived difficulty, the manipulation check, interest and the verbal and spatial ability tests before completing the posttest either immediately or they were invited back to the lab after one week to complete the posttest.

### 2.3.2 Results

**Control variables.** Means and standard deviations are reported in Table 4. Before testing the hypotheses, it was analyzed whether the six experimental groups differed in gender, age, verbal ability, spatial ability, prior knowledge, or interest. A chi-square analysis revealed no significant differences regarding gender ($p$ = .108). The remaining five control variables were analyzed by computing a MANOVA with the between-subjects factors learning strategy (drawing versus observation versus imagery) and time of testing (immediate versus delayed). It revealed no significant main effect of learning strategy, Pillai's trace = 0.076, $F$(10,332) = 1.31, $p$ = .22, partial eta² = .04, or time of testing, Pillai's trace = .042, $F$(5,165) = 1.46, $p$ = .21, partial eta² = .04, as well as no significant learning strategy x time of testing interaction, $F$ < 1. Hence, conditions were similar with regard to these entry characteristics.

**Manipulation check.** Means and standard deviations are reported in Table 4. The univariate analysis of variance (ANOVA) with learning strategy and time of testing as between-subjects factors and mental imagery rating as dependent variable revealed a significant main effect of learning strategy, $F$(1,198) = 5.53, $MSE$ = 0.73, $p$ < .01, partial eta² = .05. Post-hoc multiple comparisons revealed that participants in the imagery conditions ($M$ = 3.26, $SD$ = 0.86) reported higher values than participants in the drawing ($M$ = 2.88, $SD$ = 0.86; $p$ = .03) and in the observation conditions ($M$ = 2.79, $SD$ = 0.99; $p$ = .01), indicating that participants in the imagery conditions had followed the instructions and had engaged in mental imagery during learning. Neither the main effect of time of testing, $F$ < 1,

Table 4

*Means and standard deviations for the control variables as a function of experimental condition in Study 2*

| | Time of testing | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Immediate | | | | | | Delayed | | | | | |
| | Learning strategy | | | | | | Learning strategy | | | | | |
| | Drawing | | Observation | | Imagery | | Drawing | | Observation | | Imagery | |
| *n* | 36 | | 34 | | 34 | | 37 | | 31 | | 32 | |
| | *M* | *SD* | *M* | *SD* | *M* | *SD* | *M* | *SD* | *M* | *SD* | *M* | *SD* |
| Age | 22.47 | 2.41 | 23.24 | 2.41 | 22.76 | 2.86 | 23.92 | 3.38 | 23.16 | 3.13 | 23.31 | 2.93 |
| Verbal ability (0-20) | 14.25 | 2.43 | 13.47 | 3.06 | 13.71 | 2.33 | 13.51 | 2.87 | 14.10 | 2.98 | 13.94 | 3.17 |
| Spatial ability (-10-10) | 4.40 | 3.59 | 3.97 | 3.76 | 4.79 | 3.81 | 3.70 | 3.96 | 2.71 | 4.08 | 5.66 | 2.40 |
| Prior knowledge (0-100%) | 4.17 | 17.95 | 6.18 | 22.43 | 7.06 | 24.56 | 6.21 | 20.73 | 14.19 | 19.45 | 11.56 | 19.86 |
| Interest (1-5) | 3.52 | 0.56 | 3.71 | 1.05 | 3.38 | 0.84 | 3.57 | 1.04 | 3.69 | 0.67 | 3.49 | 0.90 |
| Mental imagery rating (0-4) | 2.71 | 0.85 | 2.84 | 0.97 | 3.21 | 0.65 | 3.05 | 0.86 | 2.75 | 1.02 | 3.30 | 0.73 |

nor the interaction, $F(2,198) = 1.08$, $MSE = 0.73$, $p = .34$, partial eta² = .01, reached significance.

**Learning outcomes.** Means and standard deviations are reported in Table 5. As in Study 1, a difference score for recognition performance was computed by subtracting the corresponding performance scores of the posttest from the prior knowledge test scores. Then, an ANOVA was conducted with learning strategy and time of testing as between factors. It revealed that the difference scores deviated significantly from zero, $F(1,198) = 96.92$, $MSE = 175.42$, $p < .001$, partial eta² = .33, indicating that participants reached significant knowledge gains between the prior knowledge test and the recognition posttest. Participants in the immediate conditions ($M = 15.14\%$, $SD = 13.86$) scored higher than participants in the delayed conditions ($M = 3.05\%$, $SD = 12.41$), $F(1,198) = 42.21$, $MSE = 175.42$, $p < .001$, partial eta² = .18. The main effect of learning strategy as well as the learning strategy x time of testing interaction were not significant, both $Fs < 1$.

For the performance in the transfer test a similar ANOVA was conducted. No significant main effects of learning strategy or time of testing and no significant interaction were found, all $Fs < 1$.

A similar ANOVA was conducted for the performance in the drawing test. It revealed a significant main effect of learning strategy, $F(2,191) = 4.36$, $MSE = 232.15$, $p < .05$, partial eta² = .04. Post-hoc multiple comparisons showed that the imagery groups ($M = 47.71\%$, $SD = 13.82$) scored lower than the observation groups ($M = 55.76\%$, $SD = 17.49$; $p = .004$) and the drawing groups ($M = 52.12\%$, $SD = 15.95$, $p = .08$). Additionally, a significant main effect of time of testing showed that overall, participants reached higher scores in the immediate ($M = 55.96\%$; $SD = 15.49$) than in the delayed posttest ($M = 47.87\%$, $SD = 15.69$), $F(1,191) = 13.91$, $MSE = 232.15$, $p < .001$, partial eta² = .07. The interaction learning strategy x time of testing was not significant, $F < 1$.

Table 5

*Means and standard deviations for the learning outcome measures and perceived difficulty as a function of experimental condition and time of testing (immediate vs. delayed) in Study 2*

| | Time of testing | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Immediate | | | | | | Delayed | | | | | |
| | Learning strategy | | | | | | Learning strategy | | | | | |
| | Drawing | | Observation | | Imagery | | Drawing | | Observation | | Imagery | |
| *n* | 36 | | 34 | | 34 | | 37 | | 31 | | 32 | |
| | *M* | *SD* | *M* | *SD* | *M* | *SD* | *M* | *SD* | *M* | *SD* | *M* | *SD* |
| Recognition (% correct) | 13.19 | 12.94 | 16.62 | 13.52 | 15.74 | 15.23 | 2.03 | 13.15 | 3.71 | 13.29 | 3.59 | 10.87 |
| Transfer (% correct) | 47.22 | 20.94 | 53.13 | 22.15 | 49.63 | 23.28 | 44.43 | 22.81 | 48.19 | 21.11 | 49.80 | 20.11 |
| Drawing (% correct) | 56.21 | 14.93 | 62.63 | 15.85 | 49.03 | 12.89 | 48.48 | 16.15 | 48.68 | 16.45 | 46.38 | 14.78 |
| Perceived difficulty (1-7) | 3.41 | 1.31 | 2.22 | 1.09 | 2.96 | 1.28 | 3.62 | 1.51 | 2.57 | 1.09 | 2.96 | 1.28 |

**Perceived difficulty.** For the perceived difficulty (see Table 5 for means and standard deviations) of the learning material, a univariate ANOVA was conducted with learning strategy and time of testing as between factors. It revealed a significant main effect of learning strategy, $F(2,198) = 13.35$, $MSE = 1.62$, $p < .001$, partial eta² = .12. Post-hoc multiple comparisons showed that the drawing groups ($M = 3.52$, $SD = 1.41$) reported higher values than the imagery groups ($M = 2.96$, $SD = 1.26$; $p = .03$), reflecting higher perceived difficulty. The latter also reported higher values than the observation groups ($M = 2.38$, $SD = 1.10$; $p = .04$). Neither the main effect of time of testing, $F(1,198) = 1.12$, $MSE = 1.62$, $p = .29$, partial eta² = .01, nor the respective interaction, $F < 1$, were significant. Perceived difficulty was negatively correlated with recognition performance ($r = -.20$, $p < .01$), transfer performance ($r = -.23$, $p < .01$) and drawing performance ($r = -.22$, $p < .01$), indicating that higher perceived difficulty was associated with lower scores on all three learning outcome measures.

**Drawing quality.** In addition to analyzing learning outcome performance, the quality of the learner-generated drawings in the drawing conditions was also analyzed (see Table 6 for means and standard deviations of drawing accuracy and drawing omissions scores). An exploratory analysis revealed that drawing accuracy scores ranged from 4.17 to 63.67% ($M = 35.19\%$, $SD = 13.17$). Means for drawing accuracy scores were significantly different from both the bottom (0%; $t(72) = 22.82$, $p < .001$) and the top of the scale (100%; $t(72) = 42.03$, $p < .001$). Drawing omissions scores ranged from 22.60 to 94.44% ($M = 54.45\%$, $SD = 13.54$). Means for drawing omissions scores were significantly different from both the bottom (0%; $t(72) = 34.69$, $p < .001$) and the top of the scale (100%; $t(72) = 28.75$, $p < .001$). Thus, participants produced drawings of medium quality.

Table 6

*Means (%) and standard deviations for drawing accuracy and omissions in the two drawing groups in Study 2*

| | Time of testing | | | |
|---|---|---|---|---|
| | Immediate | | Delayed | |
| *n* | 36 | | 37 | |
| | *M* | *SD* | *M* | *SD* |
| Drawing accuracy | 35.67 | 12.07 | 34.72 | 14.32 |
| Drawing omissions | 53.77 | 11.98 | 55.11 | 15.04 |

In line with the predictions (Hypothesis 3a), it was found that drawing accuracy was positively correlated with recognition performance ($r$ = .33, $p$ < .01), transfer performance ($r$ = .57, $p$ < .001) and performance in the drawing test ($r$ = .72, $p$ < .001), indicating that more accurate drawings were associated with higher test performance. The correlation between drawing accuracy and drawing test performance was significantly larger than the correlation between drawing accuracy and transfer performance ($z$ = 1.83, $p$ < .05) which was also significantly larger than the correlation between drawing accuracy and recognition performance ($z$ = 2.09, $p$ < .05). Drawing omissions were negatively correlated with recognition performance ($r$ = -.41, $p$ < .001), transfer performance ($r$ = -.53, $p$ < .001) and performance in the drawing test ($r$ = -.67, $p$ < .001), indicating that a higher number of main concepts that were omitted and not incorporated into the drawings was associated with lower test performance. The correlation between drawing omissions and drawing test performance was significantly larger than the correlation between drawing omissions and recognition performance ($z$ = -2.74, $p$ < .01) and marginally larger than the correlation between drawing omissions and transfer performance ($z$ = -1.59, $p$ < .10).

### 2.3.3 Summary and conclusions

The findings of the first experiment and previous studies were replicated in that benefits of drawing depended on the sensitivity of the test for these effects (van Meter & Firetto, 2013). Effects of drawing were found on a measure of higher-order knowledge, namely, visuo-spatial knowledge, but not for recognition. The findings regarding visuo-spatial knowledge suggest that externalization of a pictorial representation contributes to learning, which is why participants in the drawing and observation groups outperformed the imagery groups, supporting Hypothesis 1c. During drawing, a mental image of the learning content is constructed, similar to mental imagery, as a base for the actual drawing (van Meter & Garner, 2005). But it seems that externalizing this mental image into the real world holds an additional benefit for learners when answering questions that assess their visuo-spatial knowledge. Moreover, observing a drawing develop seems to be equally effective as self-constructing a drawing. This finding corresponds to that of Study 1, where learning from provided pictures was found to be as effective as drawing. In contrast to Study 1, however, no effects for transfer were found. The findings for transfer knowledge indicate that all three learning activities were equally effective for building a coherent mental model of the learning content.

Apparently, all three learning strategies can be considered a desirable difficulty during learning, which result in more stable and flexible knowledge (Bjork & Bjork, 2011; Bjork, 1994), as suggested by the fact that transfer scores did not decline over time. For recognition, the finding of the first experiment was replicated, that is, a significant decrease in test performance between the immediate and the delayed measurements indicating forgetting that is in line with Hypothesis 2b. However, unexpectedly, the findings regarding visuo-spatial knowledge also reflect this pattern.

A prognostic drawing effect was found in line with Hypothesis 3a and previous research (Schwamborn et al., 2010), that is, a higher number of correctly incorporated main concepts of the text into the drawings predicted better learning outcome performance. Furthermore, larger correlations of drawing accuracy and omissions were found with assessments of higher-order knowledge than with assessments of lower-order knowledge in favor of Hypothesis 3b.

The results regarding perceived difficulty of the learning material showed that participants in the drawing groups perceived the learning material as more difficult than participants in the imagery groups. The latter reported higher values than participants in the observation groups. These results were more in favor of Hypothesis 4b and suggest that neither generation of a pictorial representation nor externalization of this mentally constructed image lead to cognitive offloading.

### 2.4 Discussion

The present set of experiments investigated which factors – generation, visualization or externalization – mainly contribute to benefits of drawing and whether drawing also affects performance in a delayed posttest. In two experiments, a condition in which students were asked to generate drawings for each paragraph of an expository text during learning was compared to conditions involving learning strategies that focus on a combination of one or more of the factors generation, visualization, and externalization (see Table 1 for an overview). Study 1 focused on the comparison of the factors generation and visualization. Drawing was compared to a multimedia condition in which students learned with a combination of text and pictures, a summary condition in which students were asked to generate a written summary of the main concepts, and a text-only condition in which students were provided only with a text from which to learn. Study 2 focused on the factors generation and externalization. Drawing was compared to an observation condition in which students were provided with a text and dynamic visualizations showing a step-by-step evolution of a drawing, and a mental imagery condition in which students were asked to mentally imagine the content of each paragraph. After learning, learning outcomes were assessed with an immediate and delayed posttest. Time of testing was manipulated within subjects in Study 1 and between subjects in Study 2.

In a nutshell, previous findings were replicated that effects of drawing are sensitive to the type of assessments, that is, the two experiments showed benefits of external visualizations (self-generated or provided) on assessments of higher-order knowledge but not on assessments of lower-order knowledge. The results of the two experiments indicate that students benefited from visualization and externalization during learning and that generation even hindered learning. Drawing seemed to influence learning in a delayed test similarly to other learning strategies, that is, students showed stable knowledge or even knowledge gains for assessments of higher-order knowledge when time of testing was manipulated within subjects. When time of testing was manipulated between subjects, stable knowledge could only be identified for transfer performance.

### 2.4.1 What are the main factors that contribute to the benefits of drawing?

Based on previous research regarding task sensitivity of learner-generated drawing (Schmeck et al., 2014; van Meter & Firetto, 2013) it was predicted that benefits of drawing would occur for assessments of higher-order knowledge but not for lower-order knowledge (Hypothesis 1). In line with this assumption, Study 1 revealed positive effects of drawing and multimedia learning for transfer and visuo-spatial knowledge but not for recognition. However, not in line with this assumption, Study 2 only partially replicated this finding. It showed positive effects of drawing and observation for visuo-spatial knowledge but not for transfer knowledge or recognition. This might be due to the lack of a text-only control condition in Study 2. Mental imagery, like drawing, has been shown to be an effective strategy to increase transfer knowledge while learning from an expository text (Leopold & Mayer, 2014). Additionally, based on the results of Study 2, it seems that observing the construction of a drawing, similar to learning with multimedia, can also improve transfer knowledge. But due to the lack of a text-only control condition, this assumption could not be tested.

Furthermore, it was of interest which factors play an important role for beneficial effects of learner-generated drawing. Therefore, it was focused on three factors: generation, visualization and externalization. First, self-generating content that goes beyond what is explicitly stated in a written text is assumed to be beneficial for learning (Chi & Wylie, 2014; Foos et al., 1994). It was predicted that if generation was the main contributing factor, drawing, summarization, and mental imagery strategies which involve generation should outperform multimedia, observation and text-only strategies which do not involve generation (Hypothesis 1a). The findings do not support this assumption, since there were no comparisons for any of the learning outcome measures that favored generative learning activities; moreover, even detrimental effects of generation were found for visuo-spatial knowledge in Study 1.

Second, according to the dual coding theory learners should profit from visualizations of the learning content because they have a verbal code and an additional pictorial code of the learning content available (Clark & Paivio, 1991). If visualization was the main contributing factor, students engaging in drawing should outperform students engaging in summarizing or reading text and perform equally well as students engaging in mental imagery, multimedia learning, or observation (Hypothesis 1b). The findings of both studies mostly support this assumption as the drawing and multimedia groups outperformed the summary and text-only groups (Study 1) and no differences between experimental groups were found in Study 2 for transfer knowledge. In the latter case, all learning strategies relying on visualization performed equally well.

Third, when the visual representation of the learning content is externalized (i.e., as a learner-generated drawing) this results in a multimedia representation leading to a benefit over learning with text alone (Butcher, 2014). If externalization was the main contributing factor, students engaging in drawing should outperform students engaging in mental imagery or reading text and perform equally well as students engaging in multimedia learning, observation or summarizing (Hypothesis 1c). The results of both experiments are mostly in line with this assumption since the drawing and multimedia groups outperformed the summary and text-only groups for transfer performance and visuo-spatial knowledge in Study 1 and the drawing and observation groups outperformed the imagery group for visuo-spatial knowledge in Study 2.

Fourth, an add-on effect of all three factors was not reflected by the results of both experiments, as was predicted by the assumption that students engaging in drawing should outperform all other learning strategies (Hypothesis 1d).

To sum up, the results of both experiments indicate that visualization and externalization may be the most important contributors when learners engage in drawing during learning. This is in line with previous research showing that learner-generated and provided illustrations are equally effective for building references between verbal and visual representations of the learning content, supporting mental model construction (Hall et al., 1997; Schwamborn et al., 2011). However in Study 2, significant effects of externalization were not found for transfer, but only for visuo-spatial knowledge. This may indicate that processing an externalized visual representation is especially important in order to acquire visuo-spatial knowledge.

Unexpectedly, the results of Study 1 did not reveal beneficial effects of generation. Instead, detrimental effects of generating summaries were found on recognition. These findings suggest that engaging in text-focused generation of summaries may have led learners to focus on text-based processing; thus, fewer resources might have been left for processing visual information and, in turn, for mental model construction (Leopold & Leutner, 2012). Moreover, there was a detrimental effect of generation on visuo-spatial knowledge in Study 1. This effect could be due to learners having difficulties in extracting visuo-spatial information from the learning text, which would also explain why there was no prognostic drawing effect in this experiment.

To conclude, the present experiments suggest that benefits of drawing compared to a text-only control condition are likely due to having available a visual external representation; when learners create visualization themselves it does not seem to have any add-on benefit. These results indicate that previous findings of beneficial effects of drawing over a reading only control condition (Alesandrini, 1981; de Bock et al., 1998; Leopold & Leutner, 2012; Lesgold et al., 1977; Schmeck et al., 2014; Schwamborn et al., 2010; van Essen & Hamaker, 1990; van Meter, 2001; van Meter et al., 2006) may need to be interpreted differently. As opposed to the emphasized role of generation in the latest theoretical frameworks of learner-generated drawing (Leutner & Schmeck, 2014; van Meter & Firetto, 2013; van Meter & Garner, 2005), it seems plausible to argue that a benefit of drawing over a reading control condition can be explained in line with the dual coding theory (Clark & Paivio, 1991) and the CTML (Mayer, 2014a) in that the benefit of drawing seems to be based on learning with a verbal representation and an additional pictorial representation that is externalized. Further research is needed focusing on the importance of generation for learner-generated drawing to identify circumstances and boundary conditions under which generation of an external pictorial representation is beneficial for learning.

### 2.4.2 Are benefits of drawing more pronounced in a delayed posttest?

In previous research, effects of drawing were investigated on immediate posttests only. This study was a first attempt in extending effects of drawing to a longer time period, so a posttest with a one-week delay was included in the design. Based on research on desirable difficulties (Bjork & Bjork, 2011), it was assumed that benefits of drawing – as it can be argued that drawing is a demanding way of learning – may be more pronounced in a delayed rather than an immediate posttest. Consequently, a larger difference was expected between drawing and less demanding strategies in a delayed than in an immediate posttest for higher-order knowledge (Hypothesis 2a). The results do not support this assumption. This is likely due to the fact that time of testing was manipulated within subjects in Study 1 and participants took the same posttest twice. Thus, participants in all four experimental groups likely benefited from retrieval practice (Roediger et al., 2011) and were able to consolidate their acquired knowledge through testing. Still, manipulating time of testing between subjects in the second experiment did not reveal results in line with Hypothesis 2a. Thus, it may be unlikely that drawing can be considered a desirable difficulty. As the present study provided only a first glimpse at how drawing affects learning over a longer period, further research should investigate effects of drawing over time to broaden the body of evidence.

In line with previous research on drawing indicating a sensitivity of beneficial effects of drawing on higher-order but not on lower-order knowledge (Leutner & Schmeck, 2014; van Meter & Firetto, 2013), forgetting over time was expected regarding recognition (Hypothesis 2b). The findings in both experiments were in line with this assumption, indicating that propositional text features are stored only superficially in memory and many of these features are not remembered anymore after one week.

#### 2.4.3 Does the size of the prognostic drawing effect depend on the type of posttest?

It was of interest whether the prognostic drawing effect (Schwamborn et al., 2010), that is, higher drawing quality being associated with higher test performance, would differ depending on the type of knowledge assessed. Therefore, a positive correlation was expected between drawing quality and learning outcome performance (Hypothesis 3a) and this correlation was expected to be higher for assessments of higher-order knowledge than for assessments of lower-order knowledge (Hypothesis 3b). The results of the second experiment support this assumption revealing that students achieved higher learning outcomes the more main concepts they incorporated correctly into their drawings and the fewer main concepts they omitted from their drawings. These findings were in line with previous studies showing a positive relation between the quality of generated drawings during learning and comprehension (Schmeck et al., 2014; Schwamborn et al., 2010; van Meter, 2001; van Meter et al., 2006). Additionally, differences in the size of this relation suggest that the prognostic drawing effect is sensitive to the type of test that is administered. This means that drawing does not only benefit the acquisition of higher-order knowledge but that the quality of the constructed drawings also seems to more strongly affect the acquisition of higher-order knowledge than the acquisition of lower-order knowledge.

However, these relations were not found for drawing in Study 1. Although drawing quality in both experiments was relatively low, it seems likely that the lack of salient spatial information in the learning text used in Study 1 might be responsible for the absence of the prognostic drawing effect in this experiment.

#### 2.4.4 How does drawing relate to perceived difficulty of learning?

Based on previous research, it was hypothesized that drawing would either decrease (Hypothesis 4a) or increase perceived difficulty of the learning content (Hypothesis 4b). The findings were in line with Hypothesis 4b, supporting assumptions of Leutner et al. (2009) that instructing learners to construct free-hand drawings increases perceived difficulty because of the increased cognitive demands learners experience when concerned with the logistics and management of the drawing process. Thus, it seems that these learners experienced the task as being too difficult and, in turn, gave up on investing enough effort in meaningful learning

(Schmeck, Opfermann, van Gog, Paas, & Leutner, 2015). Providing instructional support during drawing may be used to reduce perceived difficulty, which would then allow learners to invest a substantial amount of effort in meaningful learning. This assumption was tested in Study 3.

The finding that students who engaged in mental imagery reported higher values of perceived difficulty than students who engaged in observation was somewhat unexpected, as previous research showed that mental imagery decreases perceived difficulty (Leutner et al., 2009). In that study, lower levels of perceived difficulty for mental imagery were found compared to a text-only control condition. The lack of this kind of control condition in the second experiments may explain why no similar effect of mental imagery was found. The difference between mental imagery and observation can be explained by the body of research following both cognitive theory of multimedia learning (Mayer, 2014a) and cognitive load theory (Sweller, 1999, 2005a), according to which students learn better with combinations of text and provided illustrations than with text alone because active processing during learning is sustained while it decreases extraneous mental processing.

### 2.4.5 Limitations

The results of the present research regarding transfer knowledge are to some extent inconsistent in that the first experiment revealed a positive effect of drawing and multimedia learning on transfer knowledge whereas in the second experiment, no differences were found between experimental conditions. This can be likely attributed to the comparison groups used in Study 2. The lack of a text-only control condition might be responsible for not finding a difference between conditions for transfer performance. Consistent with Study 1, no differences were found between drawing and learning with text and provided illustrations (static visualizations in Study 1 and dynamic visualizations in Study 2). Mental imagery can as well be a powerful strategy to acquire transfer knowledge (Leopold & Mayer, 2014). But when learners need to rely on the visual representation of the learning content, the findings indicate that externalization of this visual representation in the form of a drawing or provided illustrations holds a benefit over mentally imagining the content.

Moreover, the internal consistency of items used to assess learning outcome performance was especially low for the transfer and recognition test, as shown by low values of Cronbach’s alpha coefficients, which can be responsible for the somewhat inconsistent results. Thus, the construction of a test with high internal consistency might be indicated. But these low values are likely due to the fact that different concepts were measured with the items of one test to address the whole spectrum of main concepts incorporated in the learning text. Especially when the measures that are involved are not unidimensional, the alpha coefficient does not properly reflect the reliability of the scale (Schmitt, 1996). Nevertheless, it might be recommended to replicate this study with other scientific contents.

The quality of the drawings constructed during learning was relatively low, compared to other studies investigating learner-generated drawing (e.g., Schmeck et al., 2014; Schwamborn et al., 2010). Also, drawing quality influenced learning outcome performance only in one of the two experiments. This might be due to the fact that students were not sure what they were supposed to draw during learning; they may have needed support to construct more accurate drawings. There have been a few studies showing that learners need some kind of support to actually benefit from drawing (Alesandrini, 1981; Lesgold et al., 1975; van Meter et al., 2006). Support can be implemented by asking students to produce free-hand drawings and then to compare them with provided illustrations, offering the learners opportunity to identify what they left out or did not understand correctly at first (van Meter, 2001). Another way to support drawing construction is to provide learners with drawing prompts including a legend that shows all the relevant elements for the drawings and a partially pre-drawn background for constructing drawings (Schmeck et al., 2014). A replication of the present study with drawing support may be useful to address this issue and to further explore the contribution of the factors generation, visualization, and externalization to benefits of learner-generated drawing.

### 2.4.6 Conclusions

To conclude, the present study contributes to research on learner-generated drawing in that it tried to disentangle beneficial effects of drawing by considering three different contributing factors. The findings of two experiments indicate that the benefits of drawing on learning outcomes that have been shown in prior research stem mainly from the process of externalizing a visualization that drawing requires, rather than the actual generation of the drawing. Thus, these findings seem to reflect a benefit of external visualization, rather than a pure benefit of drawing. Taking the results of this study into consideration, it brings a new perspective to previously found benefits of drawing over reading a text. It could indicate that the postulated generation process may not be that essential when it comes to drawing. Moreover, it was shown that the prognostic drawing effect depends on the type of knowledge that is measured, in that the prognostic drawing effect was more pronounced for assessments of higher-order knowledge than for assessments of lower-order knowledge. Further research is especially needed on the generative component and long-term effects of drawing to clarify and specify the circumstances and boundary conditions for generation to be a beneficial contributor for drawing and for drawing to be considered a desirable difficulty for learning.

# 3 Instructional Support During Drawing: Do Different Types of Drawing Aids Affect Learning Outcomes and Perceived Difficulty Differently?

Study 3 addressed the second and third research question. In particular, the main focus was on the influence of two different types of instructional support that differ with respect to the processes being scaffolded. Moreover, it was of interest how drawing (including both unsupported free-hand drawing and also instructionally supported drawing) affects learners' perceived difficulty of learning and the quality of drawings constructed during learning.

## 3.1 Overview and Hypotheses

Learner-generated drawing is a promising cognitive learning strategy when learning from expository text. The drawing strategy can range from simple paper-pencil drawings (Gobert & Clement, 1999; van Meter et al., 2006) to the involvement of additional drawing aids that assist learners in constructing drawings like the use of cutout figures to construct drawings (Lesgold et al., 1977), or drawing prompts including illustrations of relevant objects (Schwamborn et al., 2010). These drawing aids seem to be especially helpful in assisting learners to really benefit from constructing drawings (Leutner & Schmeck, 2014; van Meter & Garner, 2005). Yet, it remains an open question of how much support during the construction of drawing is helpful for learners. Accordingly, the present study investigated how different forms of support affect learning outcome and perceived difficulty of the learning material.

To this end, a free-hand drawing condition who received no support was compared to learners receiving one of the following types of support: In the low-support condition, learners received pre-drawn backgrounds to construct drawings. This should reduce the requirement to reason about the superficial aspects of the appearance and spatial relations of irrelevant objects, thereby helping learners to fully focus on generating free-hand drawings of relevant elements. In the high-support condition, learners received different elements and labels to choose from in order to construct their drawings. This should free learners from reasoning about the visual appearance of any element, thereby enabling learners to fully focus on the spatial relations between these elements while arranging them to generate a drawing. Because distractors were introduced, learners were still required to engage in reasoning regarding the relevance of different elements. The control condition received only a text and was not instructed to engage in any drawing activity.

Two research questions guided the present study. First, it was of interest whether the two forms of instructional support during drawing that vary in how processes of creating the drawings are scaffolded would affect learning outcomes and perceived difficulty. In line with previous research and predictions by the CMDC (van Meter & Firetto, 2013), no differences were expected between conditions for assessments of lower-order knowledge (i.e., recognition, Hypothesis 1).

For assessments of higher-order knowledge (i.e., transfer and visuo-spatial knowledge) the following predictions were made in line with previous research: It was expected that all three drawing conditions would perform better in the assessments than the control condition (Hypothesis 2).

Regarding the question of how scaffolding the drawing process affects learning, different predictions can be made. One the one hand, any provision of support might aid learning irrespective of how much it constrains and guides underlying reasoning processes. In this case, the high-support and low-support conditions would be expected to outperform the no-support drawing condition to an equal extent (Hypothesis 3a). Moreover, the no-support condition would be expected to report more perceived difficulty than the other two drawing conditions, which should not differ on this measure (Hypothesis 3b). On the other hand, effects of scaffolding might depend on the type of support that is provided. First, it might be that it is most effective if learners were relieved from the need to reason about less relevant features of their drawings, while still engaging in reasoning about the visual appearance and spatial relations of relevant features. In this case, the low-support condition should score better than the high-support condition, which in turn should score higher than the no-support condition (Hypothesis 4a). Regarding perceived difficulty, the no-support condition should report higher perceived difficulty than the low-support condition, which would be expected to report a higher perceived difficulty than the high-support condition (Hypothesis 4b). Thus, the latter reduction in perceived difficulty was not expected to contribute to better learning, because it is due to a reduced cognitive engagement with the learning task. Second, scaffolding might be most effective if support is provided regarding any kind of reasoning regarding the visual appearance of drawn elements. In this case, the high-support condition should outperform the low-support condition, which should outperform the no-support condition (Hypothesis 5a). For perceived difficulty, the same pattern of results was expected as stated in Hypothesis 4b (no-support > low-support > high-support); however, the relation with learning outcomes was assumed to be different. That is, lower levels of perceived difficulty should be related to better outcomes.

Second, the relation between drawing quality and learning outcomes was examined. In line with previous research on the prognostic drawing effect (Schmeck et al., 2014; Schwamborn et al., 2010), positive correlations between drawing quality and learning outcomes were expected (Hypothesis 6). Based on assumptions of the CMDC regarding sensitivity of the drawing strategy’s benefit to well-matched posttests (van Meter & Firetto, 2013), larger positive correlations were expected between drawing quality and posttests assessing higher-order knowledge in terms of transfer and visuo-spatial knowledge than for posttests assessing lower-order knowledge in terms of recognition (Hypothesis 6a). Moreover, support was expected to also positively affect drawing quality as it helps students in constructing more complete and correct drawings. Thus, students in the low-support and high-support conditions should construct drawings of equal quality and of higher quality than the no-support condition (Hypothesis 6b).

## 3.2 Method

### 3.2.1 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. Biology, medicine, physics and sports sciences majors were excluded from the experiment. 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 no-support ($n = 39$), low-support ($n = 39$), high-support ($n = 39$) and the non-drawing control group ($n = 39$).

### 3.2.2 Materials

For the learning episode, an adapted version of the text about biomechanics in human swimming behavior used in Study 1 and Study 2 was utilized containing 983 words in 9 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 layout of the pages depended on experimental condition (see Figure 7).

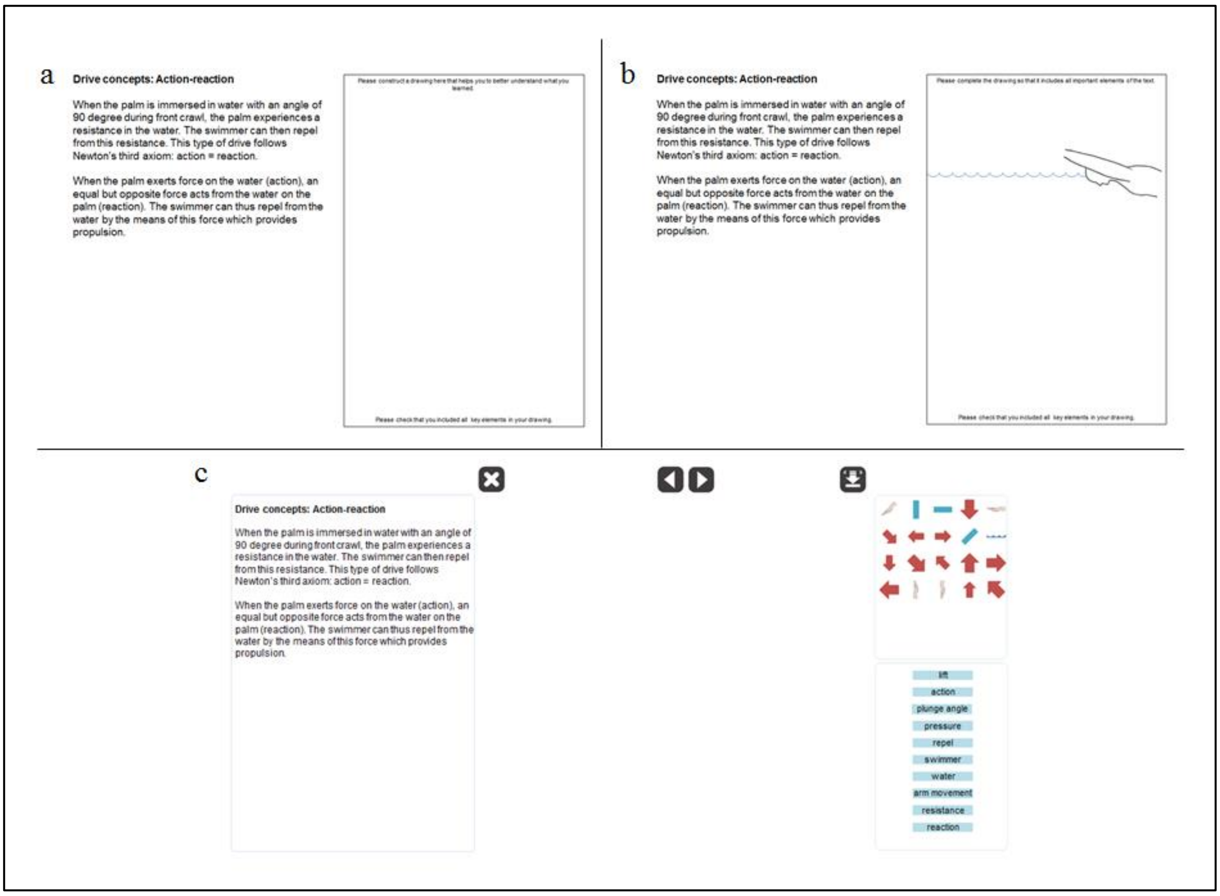


*Figure 7*. Example of a page of the learning material (translated from the German original) used in the three drawing conditions of Study 3 depending on the experimental condition: (a) no-support, (b) low-support, (c) high-support.

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 condition, the right side of each page was blank with a prompt asking participants to construct their drawing there. In the low-support condition, the right side of each page contained pre-given partial drawings and a prompt asking participants to complete these drawings. In the high-support condition, 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. Objects and labels were replaced after selection and could be used arbitrarily often. Participants received no feedback on whether they selected the correct objects.

#### 3.2.3 Measures

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

First, participants rated their 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 myofibrils."; Cronbach's alpha = .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').

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 = .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 = .65) taken from the Paper Folding Test (PFT; Ekstrom, French, & Harman, 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 = .72). The transfer test was designed to measure participants' transfer knowledge and consisted of nine multiple-choice items (Cronbach's alpha = .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 1 point for a correct answer and the resulting scores were transformed into a percentage. The drawing test consisted of four drawing items and was designed to assess participants' visuo-spatial knowledge of the learning content (Cronbach's alpha = .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 transfer and drawing tests. Inter-rater agreement was high for both the transfer test (all Krippendorff's alphas > .94), and the drawing test (all Krippendorff's alphas > .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, main idea units within each paragraph were first identified and it was assessed whether these were correctly incorporated into the drawings or not present at all. The drawing accuracy score was computed by assigning one point to participants for each correctly drawn main concept unit, yielding a maximum score of 76 points. Then, the percentage correct was calculated. For the drawing omissions score, each participant received one point for each main concept unit that was not drawn at all, yielding again a maximum score of 76 points. Similar to the drawing accuracy score, then the percentage correct was calculated. Drawing quality was only assessed for participants in the three drawing conditions because the control condition did not construct drawings during learning. Two raters scored 21 % of the data for drawing quality. Inter-rater agreement was acceptable for both drawing accuracy (all Krippendorff's alphas > .84), and drawing omissions (all Krippendorff's alphas > .90) so that one of the raters scored the remaining data.

Third, perceived difficulty was measured using three items (e.g., "How easy or difficult was it for you to learn something about biomechanics in human swimming behavior?"; Paas, 1992) that were rated on a 7-point Likert-scale ranging from *very easy* to *very difficult*.

### 3.2.4 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. The presentation of the measures, except for the drawing test, was web-based. The presentation of the learning material in the no-support, low-support and control 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 high-support condition, a Windows-based tablet-computer (Microsoft Surface Pro 2 with an integrated 10.6'' touchscreen) was used because the application that was used to implement the learning material in the high-support condition did not run on an iOS-based tablet-computer. The learning environment of the high-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.

### 3.2.5 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-computer. 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 they should engage in has proven to be effective for learning. Additionally, participants in the drawing conditions (no-support, low-support and high-support conditions) were instructed to thoroughly read the learning text 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]. During the learning episode, participants in the no-support and low-support conditions were able to use all previously mentioned functions of Notability, whereas participants in the control condition were only able to scroll through the learning material. 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.

[1] Participants in the high-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.

## 3.3 Results

### 3.3.1 Control variables

Means and standard deviations of the control variables are reported in Table 7. Before testing the hypotheses, it was 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 ($p$ = .88). For the remaining control variables, a multivariate analysis of variance (MANOVA) was computed 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, $p$ = .21, partial eta$^2$ = .05. Thus, the experimental conditions were similar with respect to participants' entry characteristics.

### 3.3.2 Learning outcomes

Means and standard deviations are reported in Table 8. To test the assumption that a drawing strategy does not affect assessments of lower-order knowledge, a repeated measures analysis of variance was conducted with the scores in the prior knowledge test and recognition posttest as dependent variables, time of testing (prior knowledge versus posttest) as within-subjects factor and experimental condition as the between-subjects factor. It revealed a significant main effect of time of testing, $F$(1,152) = 776.89, $MSE$ = 181.90, $p < .001$, partial eta$^2$ = .84, reflecting that participants reached substantial knowledge gains between the assessment of prior knowledge and the recognition posttest. The main effect of experimental condition, $F$(3,152) = 1.45, $MSE$ = 211.91, $p$ = .23, partial eta$^2$ = .03, as well as the respective interaction were both not significant, $F$(3,152) = 1.81, $MSE$ = 181.90, $p$ = .15, partial eta$^2$ = .04.

Table 7

*Mean values and standard deviations of the control variables age, subjective ability to draw, enjoyment regarding drawing, prior knowledge, verbal ability and spatial ability for the four experimental groups in Study 3*

| Group | Type of control variables | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | *n* | Age | | Subjective ability to draw (1-7) | | Enjoyment regarding drawing (1-7) | | Prior knowledge (0-100%) | | Verbal ability (0-20) | | Spatial ability (-10-10) | |
| | | *M* | *SD* | *M* | *SD* | *M* | *SD* | *M* | *SD* | *M* | *SD* | *M* | *SD* |
| No-support | 39 | 20.85 | 2.69 | 3.36 | 1.44 | 4.49 | 1.75 | 26.28 | 13.56 | 13.56 | 2.53 | 5.00 | 4.00 |
| Low-support | 39 | 22.67 | 3.58 | 3.77 | 1.77 | 4.72 | 1.99 | 27.56 | 16.22 | 13.21 | 3.23 | 5.05 | 2.74 |
| High-support | 39 | 22.10 | 2.68 | 2.77 | 1.42 | 3.77 | 1.61 | 27.68 | 12.87 | 13.46 | 3.05 | 4.85 | 3.08 |
| Control | 39 | 20.69 | 2.26 | 3.51 | 1.59 | 4.59 | 1.67 | 27.05 | 16.25 | 13.85 | 3.22 | 5.33 | 3.16 |

Table 8

*Means and standard deviations for the learning outcome measures and perceived difficulty as a function of the four experimental groups in Study 3*

| Group | | Type of test | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | *n* | Recognition (% correct) | | Transfer test (% correct) | | Drawing test (% correct) | | Perceived difficulty (0-6) | |
| | | *M* | *SD* | *M* | *SD* | *M* | *SD* | *M* | *SD* |
| No-support | 39 | 68.08 | 12.33 | 51.85 | 18.96 | 47.54 | 22.83 | 2.71 | 1.20 |
| Low-support | 39 | 71.28 | 13.11 | 46.15 | 19.17 | 49.17 | 19.26 | 3.26 | 1.08 |
| High-support | 39 | 65.13 | 14.35 | 51.25 | 20.96 | 49.52 | 17.22 | 3.03 | 1.06 |
| Control | 39 | 74.36 | 12.99 | 41.03 | 17.69 | 37.96 | 20.10 | 2.37 | 1.28 |

To test the predictions regarding the question of how support during drawing affects higher-order learning, a univariate analysis of variance (ANOVA) was conducted with experimental condition as between-subjects factor and transfer performance as dependent variable. Experimental conditions differed significantly, $F(3,152) = 2.71$, $MSE = 369.79$, $p < .05$, partial $\eta^2 = .05$. Post-hoc multiple comparisons showed that both the high-support ($p = .02$) and the no-support conditions ($p = .01$) outperformed the control condition. The high-support and no-support conditions did not differ significantly ($p = .90$). There were no differences between the low-support and the other three conditions (all $p_s > .19$).

A reduced sample of participants ($N = 153$) was used for the analysis of drawing test performance, because four participants terminated the study prematurely after the transfer test and did not work on the drawing test. An ANOVA revealed that the experimental conditions differed significantly, $F(3,149) = 2.89$, $MSE = 398.33$, $p < .05$, partial $\eta^2 = .06$. The high-support ($p = .01$), low-support ($p = .02$) and no-support conditions ($p = .04$) scored better than the control condition, thereby performing equally well (all $p_s > .66$).

### 3.3.3 Perceived difficulty

Means and standard deviations are reported in Table 8. To test the assumptions regarding the question of how drawing aids affect perceived difficulty of the learning material, an ANOVA revealed that the experimental conditions differed with regard to perceived difficulty ratings, $F(3,152) = 4.33$, $MSE = 1.34$, $p < .01$, partial $\eta^2 = .08$. Post-hoc multiple-comparisons showed that the low-support condition reported higher values of perceived difficulty than the no-support ($p = .04$) and the control conditions ($p = .001$). Moreover, the high-support condition reported higher values than the control condition ($p = .01$). Perceived difficulty was negatively correlated with recognition performance ($r = -.28$, $p < .001$) indicating that higher perceived difficulty was associated with poorer recognition test performance, but was not significantly correlated with transfer and drawing test performance (both $|r_s| < .11$ , $p_s > .21$).

### 3.3.4 Drawing quality

The quality of the learner-generated drawings constructed in the three drawing conditions was analyzed to test Hypotheses 6, 6a and 6b. An exploratory analysis showed that drawing accuracy scores ranged from 3.33% to 71.72% ($M$ = 28.66%, $SD$ = 14.15). Means for drawing accuracy were significantly different from the bottom (0%; $t(116) = 21.92$, $p < .001$) and the top of the scale (100%; $t(116) = 54.55$, $p < .001$). Drawing omissions scores ranged from 18.99% to 92.66% ($M$ = 61.13%, $SD$ = 15.72). Means for drawing omissions scores were significantly different from the bottom (0%; $t(116) = 42.05$, $p < .001$) and the top of the scale (100%; $t(116) = 26.74$, $p < .001$). Thus, participants produced drawings of medium quality. There were no floor or ceiling effects. Means and standard deviations of drawing accuracy and drawing omissions scores are reported in Table 9.

Table 9

*Means (%) and standard deviations of drawing accuracy and omissions for the three drawing groups in Study 3*

| Group | Drawing quality | | | | |
|---|---|---|---|---|---|
| | *n* | Drawing accuracy | | Drawing omissions | |
| | | *M* | *SD* | *M* | *SD* |
| No-support | 39 | 36.00 | 11.80 | 51.84 | 12.38 |
| Low-support | 39 | 33.05 | 13.83 | 56.92 | 14.88 |
| High-support | 39 | 16.93 | 8.00 | 74.63 | 9.29 |

In line with the prognostic drawing effect (Schwamborn et al., 2010), drawing accuracy was associated with higher performance in the recognition ($r = .42$, $p < .001$), transfer ($r = .21$, $p < .05$), and drawing test ($r = .28$, $p < .01$). Unexpectedly, the correlation between drawing accuracy and recognition performance was significantly larger than the correlation between drawing accuracy and transfer performance ($z = 1.98$, $p < .05$), and marginally larger than the correlation between drawing accuracy and drawing test performance ($z$ = 1.33, $p$ < .10). Drawing omissions were negatively correlated with recognition ($r$ = -.36, $p$ < .001) and drawing test performance ($r$ = -.24, $p$ < .05) reflecting that a higher number of concepts that were omitted from drawings was associated with lower test performance. Unexpectedly, drawing omissions were not significantly correlated with transfer test performance ($r$ = -.08, $p$ = .42). Surprisingly, the correlation between drawing omissions and transfer performance was significantly smaller than the correlation between drawing omissions and recognition performance ($z$ = 2.55, $p$ < .01), and also marginally smaller than the correlation between correlation between drawing omissions and drawing test performance ($z$ = 1.41, $p$ < .10). The latter two correlations did not differ significantly ($z$ = -1.11, $p$ = .13).

A MANOVA was computed with drawing accuracy and drawing omissions as dependent variables and experimental condition as between-subjects factor to test the assumption that support aids learners to construct more accurate and complete drawings. It revealed that drawing quality differed significantly between conditions, Pillai's trace = 0.406, $F$(4,228) = 14.52, $p$ < .001, partial eta² = .20. Consecutively conducted ANOVAs showed that the drawing conditions differed significantly on both drawing accuracy, $F$(2,114) = 31.26, $MSE$ = 131.49, $p$ < .001, partial eta² = .35, and drawing omissions. $F$(2,114) = 36.33, $MSE$ = 153.65, $p$ < .001, partial eta² = .39. Bonferroni adjusted multiple comparisons revealed that the high-support condition constructed less accurate drawings than the low-support ($p$ < .001) and the no-support conditions ($p$ < .001). Moreover, the high-support condition omitted more main concepts from their drawings than the low-support ($p$ < .001) and the no-support conditions ($p$ < .001).

## 3.4 Discussion

Study 3 investigated how different types of drawing support, which differed with respect to how the drawing process is scaffolded, affect learning outcomes and perceived difficulty of the learning material. To this end, a free-hand drawing condition that received no additional instructional support was compared with two drawing conditions comprising instructional support and a non-drawing control condition. The low-support condition received pre-drawn backgrounds in order to reduce the need to reason about superficial aspects regarding the appearance and spatial relations of irrelevant objects, thereby enabling learners to focus on generating free-hand drawings of relevant elements. The high-support condition received different elements and labels to choose from in order to construct drawings. This should free learners from reasoning about the visual appearance of any element of the drawings, thereby helping learners to fully focus on the spatial relations between these elements.

In brief, findings of previous research were replicated in that learner-generated drawing benefitted the acquisition of higher-order but not lower-order knowledge. Moreover, a higher quality of constructed drawings was associated with better test performance. None of the implemented types of support further promoted learning. On the contrary, baseline support even hindered learning of transfer knowledge. Also, support did not reduce perceived difficulty of the learning material as had been expected

### 3.4.1 Which type of support do learners need during drawing?

Previous research on drawing indicates that some kind of support during drawing is necessary in order for learners to benefit more from a drawing strategy (Alesandrini, 1981; Lesgold et al., 1977, 1975; Schmeck et al., 2014; Schwamborn et al., 2010; van Meter, 2001; van Meter et al., 2006). In Study 3, it was of interest whether the beneficial effect of instructional support depends on the types of processes that are scaffolded. Unexpectedly, the results showed that neither the low-support nor the high-support conditions outperformed the no-support drawing condition. The low-support condition even performed equally well as the text-only control condition on the transfer test. The results indicate that none of the present drawing aids helped learners to benefit more from drawing, thereby contradicting previous research on instructional support during drawing (for overviews see Leutner & Schmeck, 2014; van Meter & Firetto, 2013; van Meter & Garner, 2005).

Concerning the low-support condition, it may have been that the visual representations used in the pre-drawn backgrounds were not suitable for supporting the drawing process. The pre-drawn backgrounds were designed to constrain the drawing process for learners in order to reduce reasoning about irrelevant features of the drawings. Potentially, the visual representations used in the low-support condition interfered with the learners' own mental representations built during initial reading of the text. Thus, additional reasoning may have been necessary to construct drawings of the learning content that matched the pre-drawn backgrounds rather than the learners' own conceptions of how elements should look like. The need to overcome these initial conceptions may have increased rather than reduced the difficulties of generating a drawing as indicated by higher levels of perceived difficulty reported in the low-support condition. In sum, provided pre-drawn backgrounds were unsuited to increase drawing quality and learning outcomes beyond a beneficial effect of unsupported drawing.

Also unexpectedly, the high-support condition did not outperform the no-support condition in the learning outcome test. In the high-support condition, meaningful elements as well as distractors were presented from which learners could choose in order to generate drawings by spatially arranging these elements. While this approach was found to be non-beneficial for first-graders during prose learning (Lesgold et al., 1975), it was expected of adult learners to benefit from this type of support because reasoning about the relevance of elements should facilitate selecting processes that are important for learning with learner-generated as well as provided visual representations (Mayer, 2014a; van Meter & Firetto, 2013). However, this assumption was neither reflected in learning outcome nor drawing quality. The high-support condition performed equally well as the no-support condition with respect to transfer and visuo-spatial knowledge, both outperforming the control condition. Moreover, the high-support condition constructed drawings of least quality compared with the other two drawing conditions. Finally, they reported more perceived difficulty compared to the control condition indicating that reasoning about relevant and irrelevant elements was not facilitated. Although they constructed drawings of less quality, learners in the high-support condition were able to benefit from drawing to the same extent as learners who received no drawing support. Similar to Schmeck et al. (2011), a toolbar including all relevant pictorial elements for constructing drawings was also used in the high-support condition of the present study. Moreover, they provided learners with a partly pre-drawn background for drawing construction. In contrast to this approach, pre-drawn backgrounds were not included in the high-support condition but distractors were added to the pool of pictorial elements. Taken together, this finding may indicate that the distractors hindered learning also in adults but that this type of support may hold great potential for constraining the drawing process if the distractors were omitted (Lesgold et al., 1977, 1975; Schwamborn et al., 2011).

### 3.4.2 Can support during drawing reduce perceived difficulty?

Based on previous research on drawing aids, it was predicted that providing instructional support during drawing should reduce perceived difficulty of the learning materials (Schmeck et al., 2014; Schwamborn et al., 2011). Unexpectedly, neither providing learners with pre-drawn backgrounds (low-support condition), nor providing learners with different elements to construct drawings (high-support condition) reduced perceived difficulty; both conditions reported more perceived difficulty compared with the control condition. Pre-drawn backgrounds even increased perceived difficulty compared to unsupported drawing. Reducing the need to reason about irrelevant parts of a drawing by providing learners with pre-drawn backgrounds seemed to be not enough of a constraint for learners to benefit from this type of support. The increased values of perceived difficulty indicate that learners invested a significant amount of cognitive resources in managing the drawing process (Schmeck et al., 2014; Schwamborn et al., 2011) which, in turn, could not be invested in additional meaningful learning. This finding is in line with the assumption that learners' own conceptions built during initial reading interfered with the pictorial representations used in the pre-drawn backgrounds. Thus, additional reasoning to overcome these conceptions was necessary reflected by higher levels of perceived difficulty.

Surprisingly, unsupported drawing did not increase perceived difficulty as would have been expected from previous research (e.g., Leutner et al., 2009). However, this finding could depend on the learning material that was used in the present study. Whereas Leutner et al. (2009) used a learning material that involved rather abstract visual representations of molecules the present study used a learning material including processes that predominantly involved objects that are visible to the human eye. Therefore, learners could more heavily rely on their prior knowledge regarding the visual appearance of these objects (e.g., the human body or a hand). Thus, participants were not negatively affected by unsupported drawing with respect to perceived difficulty.

### 3.4.3 Prognostic drawing effect

It was of interest whether the type of posttest also affects the prognostic drawing effect, which describes the finding that 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). In line with the prognostic drawing effect, a positive correlation between drawing quality and test performance was expected (Hypothesis 6); moreover, this correlation was expected to be more pronounced for measures of higher-order knowledge than of lower-order knowledge (Hypothesis 6a). The results replicate previous findings regarding the prognostic drawing effect, revealing a positive relation between drawing accuracy and learning outcome performance. Also in favor of Hypothesis 6, it was found that a higher number concepts being omitted in the drawing was associated with lower recognition and drawing test performance. Unexpectedly, the same relation did not hold true for transfer performance. There was no support for Hypothesis 6a, because the correlation for lower-order knowledge was not smaller but even higher than the correlation for assessments of higher-order knowledge. These results might be influenced by the way how drawing quality was determined. Drawing quality was assessed as the number of main ideas from the text correctly translated into a drawing; hence, drawing quality more strongly resembles what is measured in the recognition and visuo-spatial knowledge test. On the other hand, answering transfer questions involves learners to draw inferences that go beyond what is explicitly stated in the text, thereby showing less resemblance with what was coded to obtain a measure of drawing quality.

Furthermore, it was expected that both types of drawing support should positively influence drawing quality because it should assist learners in constructing more complete and correct drawings (Hypothesis 6b). There was no support for this hypothesis as the low-support and no-support conditions did not differ in drawing quality; the high-support condition even produced drawings of lesser quality than the other two drawing conditions. That is, instructional support did not help learners to construct more accurate drawings. It seems that the pre-drawn backgrounds used in the low-support condition did not provide enough constraints to help learners construct higher quality drawings. Furthermore, instructing learners to select elements from various options to construct drawings including distractors even led to learner-generated drawings of lower quality, resulting from incorporating distractors into the drawings. Nevertheless, this type of support yielded similar results regarding learning outcome performance and increased learning compared to a non-drawing control group. 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 & 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 free-hand drawing.

### 3.4.4 Limitations

The results regarding transfer knowledge might have been confounded with the answer format that was used within the items to assess transfer performance. In Study 3, multiple-choice items were used to assess transfer performance. It might be recommended to use open-ended questions instead of multiple-choice items that better capture learners’ deep understanding of the learning content. When asked to generate a more elaborate answer to a transfer question, learners have the opportunity to explain their understanding more thoroughly. Thus, learners’ ability to transfer the acquired knowledge to new contexts might have been underestimated.

In contrast to the predictions, drawing quality was not positively influenced by instructional support. The lack of finding an improvement in drawing quality might be due to the way drawing quality was assessed. Typically, drawing quality is assessed by counting the number of main idea unit that are correctly incorporated into the drawings (Schmeck et al., 2014; Schwamborn et al., 2010; van Meter, 2001; van Meter et al., 2006). This approach requires the design of a coding scheme that defines how a correctly drawn main idea unit looks like. Accordingly, such a coding scheme is not able to capture the quality of a drawing correctly when learners use other ways to represent the learning content than the intended ones. Especially in the high-support condition where learners frequently used the distractor items to construct drawings, this explanation could be important in that it was not able to fully capture learners’ understanding of the content in the drawing quality ratings. Thus, it might be recommended to use think-aloud protocols in addition to drawing quality ratings in order to gain a better understanding of what learners understand during learning and how this understanding is reflected in their drawings.

### 3.4.5 Conclusions

To sum up, Study 3 contributes to research on learner-generated drawing in that it investigated which types of instructional support are helpful for learners to benefit more from constructing drawings during learning. The findings of the present study indicate that neither reducing the requirement to reason about irrelevant elements of the drawings alone, nor reducing the need to reason about the visual appearance of any element 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 the drawings being constructed free-hand or with additional instructional support. Future research is needed to investigate under which circumstances additional drawing aids can increase benefits of learner-generated drawing and how drawing aids have to be designed in order to be beneficial for learning.

# 4 General Discussion

Learning with external visual representations is very common in formal and informal educational settings. Multimedia material (a combination of text and pictures) has shown to be effective for learning. Following the assumptions of Mayer's CTML (Mayer 2009, 2014a) and Schnotz' ITPC (Schnotz, 2014; Schnotz & Bannert, 2003), the integration of verbal and pictorial information into a coherent mental model of the learning content is crucial for benefits of multimedia learning. However, learners tend to ignore provided pictures and show only few attempts to integrate (Hegarty & Just, 1993; Mason et al., 2013, 2015; Scheiter & Eitel, 2015). To provide learners with a drawing strategy instruction while learning can foster integration of verbal and pictorial information (van Meter & Firetto, 2013; van Meter & Garner, 2005).

During drawing learners construct an external pictorial representation that depicts the most important elements and their spatial relations described in a written text (Schwamborn et al., 2010). Learner-generated drawing has been shown to be effective for learning in different content domains and age groups (for overviews see Ainsworth et al., 2011; van Meter & Firetto, 2013). The present thesis addressed three research questions: First, it aimed to investigate which of three factors – generation, visualization, and externalization – mainly contributes to positive effects of learner-generated drawing. Second, it examined how several boundary conditions would affect benefits of drawing – namely, the type of posttest, the quality of drawings constructed during learning, test delay as well as type of instructional support. Third, it was of interest how a drawing strategy instruction influences learners' perceived difficulty of learning. To this end, three empirical studies were conducted that addressed the three research questions.

In the following, the main results of the three studies conducted within the present thesis will be summarized. Then, the results will be discussed against the theoretical background and with regard to theoretical implications. Subsequently, practical implications will be derived and strengths of the present thesis will be outlined. Next, limitations and future directions are discussed before the general discussion will be completed with a concluding statement.

## 4.1 Summary of Main Results

Studies 1 and 2 investigated which factors mainly contribute to benefits of drawing and whether drawing also affects performance in a delayed posttest. In both studies, a condition in which students were asked to generate drawings for each paragraph of an expository text during learning was compared with conditions involving one of several learning strategies that comprised a combination of one or more of the factors generation, visualization, and externalization (see Table 1 for an overview).

Study 1 focused on the comparison of the factors generation and visualization. Drawing was compared with a multimedia condition in which students learned with a combination of text and pictures, a summary condition in which students were asked to generate a written summary of the main concepts, and a text-only condition. After learning, learning outcomes were assessed with an immediate and delayed posttest, whereby the time of assessment was manipulated within subjects. No group differences were found for lower-order knowledge replicating previous findings that benefits of drawing are more likely emerge on assessments of higher-order knowledge that match the coherent mental model constructed during drawing (Leutner & Schmeck, 2014; van Meter & Firetto, 2013; van Meter & Garner, 2005). The results for higher-order knowledge showed that learning with external pictorial representations (self-generated or provided) increased learning indicating that visualization is the factor contributing most to benefits of drawing. Unlike previous research on the prognostic drawing effect (Schmeck et al., 2014; Schwamborn et al., 2010), the quality of drawings constructed during learning did not predict learning outcome. Because the results regarding time of testing – test scores being higher in the delayed than in the immediate assessment of higher-order knowledge – likely were confounded with benefits of retrieval practice (Roediger & Karpicke, 2006), time of testing was included as a between-subjects factor in Study 2.

Study 2 focused on the factors generation and externalization. Drawing was compared with an observation condition in which students were provided with a text and dynamic visualizations showing a step-by-step evolution of a drawing, and a mental imagery condition. After learning, learning outcomes were assessed with an immediate and delayed posttest that was manipulated between subjects. The results of Study 1 were replicated regarding the influence of type of posttest and the influence of an external pictorial representation being the main contributing factor to benefits of drawing. Findings regarding time of testing indicate that drawing does not constitute a desirable difficulty (Bjork & Bjork, 2011; Bjork, 1994). Drawing quality was positively associated with learning in line with the prognostic drawing effect (Schmeck et al., 2014; Schwamborn et al., 2010). As expected, the size of the prognostic drawing effect also depended on the type of posttest in that it was larger for assessments of higher-order knowledge than for assessments of lower-order knowledge. Moreover, free-hand drawing increased perceived difficulty indicating that drawing left learners with less cognitive resources available to engage in meaningful learning.

Study 3 focused on which type of instructional support is most effective in order to help learners benefit more from constructing drawings. Therefore, a no-support drawing condition was contrasted to a low-support condition, a high-support condition, and a text-only control condition. In the low-support condition, learners received drawing instructions and pre-drawn backgrounds for drawing construction, so that the requirement to reason about the appearance and spatial relations of less relevant elements was reduced. In the high-support condition, learners received drawing instructions and a box containing different elements and labels to use for drawing construction that should reduce unnecessary reasoning about the visual appearance of any of the drawings' elements, thereby enabling learners to fully focus on the spatial relations between elements during drawing. Learning outcome was assessed immediately after learning. Instructional support did not increase benefits of drawing; moreover, it increased rather than reduced perceived difficulty. Drawing with and without instructional support fostered higher-order knowledge learning to an equal extent. The results indicate that constructing drawings that matched the pre-drawn backgrounds rather than learners' initial conceptions built during reading may have increased rather than decreased the difficulties of drawing construction. Also, the findings indicate that the distractor elements used in the high-support condition caused this support measure to be ineffective. Drawing quality was positively associated with learning outcome; however, the prognostic drawing effect was not larger for assessments of higher-order knowledge than for assessments of lower-order knowledge.

## 4.2 Theoretical Implications

In the following, the results of three studies will be discussed against the backdrop of the theoretical background introduced in Chapter 1. First, the influence of main contributing factors to benefits of drawing will be targeted, before possible theoretical implications with regard to boundary conditions that influence the effectiveness of drawing and the influence of drawing on perceived difficulty will be discussed.

### 4.2.1 What are the main contributing factors to learner-generated drawing?

Within Study 1 and Study 2, the influence of three factors – generation, visualization and externalization – on positive effects of drawing was investigated. Against the backdrop of the CMDC (van Meter & Firetto, 2013), these factors can be associated with different processes involved in the construction of the mental representations during drawing. First, generation can be associated with the construction of all four mental representations built during drawing (surface representation, propositional network, mental model, and perceptual image) because the learner is viewed as an active agent. Similar to the CTML (Mayer, 2009, 2014a) the learner actively engages in selecting relevant elements in the text to build a surface representation that is then organized into the propositional representation. The construction of the mental model also involves generative processes in terms of deriving a visuo-spatial representation, integrating verbal and pictorial information, and applying prior knowledge. Moreover, generation plays an important role when a perceptual image is derived that serves as a base for the construction of the actual drawing. Second, visualization in terms of dual coding takes place when visuo-spatial information is derived from the text-based propositional representation during mental model construction. Furthermore, visualization is involved when visuo-spatial information is extracted from the mental model in deriving a perceptual image. Lastly, externalization is important during actual drawing; when the mental perceptual image is externalized as a drawing into the real world.

The results of Study 1 and Study 2 indicate that the benefits of learner-generated drawing on learning outcomes found in previous research (for overviews see van Meter & Firetto, 2014; van Meter & Garner, 2005) stem mainly from externalizing an additional visual representation of the learning content that drawing requires rather than the actual generation of the drawing. These results can bring a new perspective to previously found benefits of drawing over reading a text (e.g., van Meter, 2001; van Meter et al., 2006) as they indicate that the postulated generation process may not be that crucial when it comes to drawing. This is in line with previous research showing that provided and learner-generated visualizations both equally support making inferences between the verbal and the pictorial representations during mental model construction (Hall et al., 1997; Schwamborn et al., 2011). However in Study 2, effects of externalization were found only for visuo-spatial knowledge but not for transfer performance. This finding indicates that processing an external pictorial representation of the learning content may be especially important when it comes to the acquisition of visuo-spatial knowledge.

Unexpectedly, Study 1 even showed detrimental effects of generation. In particular, generating summaries had negative effects on recognition performance. This suggests that text-focused generation of verbal summaries may have triggered learners to concentrate on text-based processing. Therefore, fewer resources may have been left for processing visuo-spatial information and, in turn, for constructing a coherent mental model (Leopold & Leutner, 2012). Moreover, writing summaries has been shown to produce mixed effects with some studies showing positive effects while others did not (for an overview see Dunlosky et al., 2013). It seems that extensive training is necessary in order for learners to really benefit from writing summaries (Dunlosky et al., 2013). Thus, drawing may be compared to other text-focused generative strategies that require less training such as note-taking or self-explaining.

To sum up, the results of Study 1 and Study 2 suggest that benefits of drawing over reading a text likely stem from the availability of an external visual representation of the to-be-learned content, whereas self-generating these visual representations does not seem to have any additional benefit. Against the backdrop of these results, one can argue that multimedia learning and drawing involve similar generation processes. Both, the CMDC and the CTML view learners as active agents who are actively involved in constructing mental representations of the learning content, and building connections between these representations as well as prior knowledge (Mayer, 2009, 2014a; van Meter & Firetto, 2013; van Meter & Garner, 2005). Thus, the results may indicate that previously found benefits of drawing over a text-only control condition (Alesandrini, 1981; de Bock, Verschaffel, & Janssens, 1998; Leopold & Leutner, 2012; Lesgold et al., 1977; Leutner & Schmeck, 2014; Schwamborn et al., 2010; van Essen & Hamaker, 1990; van Meter, 2001; van Meter et al.,

2006) should be interpreted differently. In contrast to the emphasized role of generation in the latest theoretical frameworks on drawing (Leutner & Schmeck, 2014; van Meter & Firetto, 2013; van Meter & Garner, 2005), it seems plausible to argue that benefits of drawing over a text-only control condition can be explained in line with the dual coding theory (Clark & Paivio, 1991; Paivio, 2006) and the CTML (Mayer 2009, 2014a). That is, positive effects of drawing seem to be based on learning with a verbal representation and an additional pictorial representation that is externalized. One can argue that learner-generated drawing supports multimedia learning in that it forces learners to engage in integrating verbal and pictorial information because pictorial information has to be processed in order to derive the visuo-spatial information from the verbal information of the text. Integration has been shown to be crucial in order for learners to benefit from a multimedia message (Hegarty & Just, 1993; Mason et al., 2015; Scheiter & Eitel, 2015). It might be that self-generating the pictorial representation does not provide an additional benefit when it comes to drawing because multimedia learning already involves generative processes in that learners actively select, organize and integrate information to build a coherent mental model. Further research should address the importance of generation for learner-generated drawing to identify circumstances and boundary conditions under which the process of self-generating an external visual representation is helpful for learning.

### 4.2.2 Which boundary conditions influence benefits of drawing?

The following paragraph relates to the second research question on boundary conditions influencing benefits of drawing, and is divided into four parts: the influence of the type of posttest on (1) benefits of drawing and (2) on the size of the prognostic drawing effect, (3) test delay, and (4) instructional support.

First, the assumption was tested that positive effects of drawing are more likely to be detected on assessments of higher-order knowledge than on assessments of lower-order knowledge (Leutner & Schmeck, 2014; van Meter & Firetto, 2013; van Meter & Garner, 2005). The results of all three studies support this assumption as benefits of a drawing strategy instruction were found only for higher-order assessments of transfer or visuo-spatial knowledge but not for lower-order assessments of recognition performance. Thus, the present studies provide further evidence that well-matched posttests are important in order to be able to reveal benefits of learner-generated drawing. It is more likely to find positive effects of drawing if the posttest aligns with the characteristic of learners' mental knowledge representation constructed during drawing (van Meter & Firetto, 2013). For example, the mental model constructed during drawing provides a functional analog system of what is described in the text. This allows the learner to reason about what would happen to the system if certain parts were manipulated in a way of a cause-and-effect chain. Transfer tests require exactly this type of reasoning. Thus, this test matches the characteristics of the knowledge representation. On the other hand, the mental model does not include verbal details from the instructional material (Schnotz, 2002), which can explain why a drawing strategy instruction fails to improve recognition performance. The present thesis contributes to research on learner-generated drawing as it further broadens the body of research showing benefits of drawing for assessments of visuo-spatial knowledge. Because learners who engage in drawing are more likely to build a knowledge representation that contains visuo-spatial information, it is not surprising that learners who construct drawings during learning are better able to reproduce these drawings in a posttest (Leopold & Leutner, 2012; Leutner et al., 2009; Schwamborn et al., 2011). In the present studies, however, learners were not required to reproduce the drawings of the learning episode in the drawing test; rather, they had to apply their acquired knowledge to new contexts. The results provide further evidence for the assumption that learners build a depictive mental model during learning, which is then translated into a perceptual image to generate the drawing (van Meter & Firetto, 2013). Thus, drawing improves performance on a visuo-spatial knowledge test because this measure matches the knowledge represented in the mental model and perceptual image.

Second, it was examined whether the size of the prognostic drawing effect (Schwamborn et al., 2010) also depends on the type of posttest. That is, the positive correlation between the quality of drawings constructed during learning and learning outcome performance was expected to be higher for assessments of higher-order knowledge than for assessments of lower-order knowledge. The results of Study 2 were in line with this assumption. This finding is not surprising if one considers that the characteristics of the mental model and perceptual image built during drawing nicely align with the tasks of a transfer or visuo-spatial knowledge test (van Meter & Garner, 2013). The lack of replicating this finding in Study 3 is most likely due to the fact that transfer knowledge was assessed with a multiple-choice test rather than with open-ended questions which were used to assess transfer performance in Study 1 and Study 2. Thus, learners were constrained with respect to the possibility to explain their knowledge thoroughly. One can argue that the results of Study 2 could have been replicated, if open-ended questions had been used to assess transfer performance. More research is required to substantiate the finding that the size of the prognostic drawing effect is larger for assessments of higher-order knowledge than for assessments of lower-order knowledge. In particular, it is recommended to replicate the findings using different types of higher-order knowledge posttests such as problem-solving (Arcavi, 2003; Rellensmann et al., 2016; van Essen & Hamaker, 1990; van Meter et al., 2006) or comprehension tests (Alesandrini, 1981). Moreover, open-ended questions should be used to be able to gain a comprehensive insight into a learner’s knowledge representation built during drawing.

Third, it was investigated whether drawing constitutes a desirable difficulty (Bjork & Bjork, 2011; Bjork, 1994). The gap between drawing and less demanding learning strategies was expected to be larger in a delayed rather than an immediate posttest of transfer knowledge. The results of Study 1 and Study 2 did not support this assumption. The differences between drawing and other learning strategies were independent of time of testing. The results of Study 1 indicate that learners who received drawing strategy instructions benefited from retrieval practice (Roediger & Karpicke, 2006; Roediger et al., 2011) to the same extent as learners engaging in other learning strategies as they were able to consolidate their acquired knowledge through testing. When controlling for effects being confounded with retrieval practice in Study 2, the results indicate that following a drawing strategy instruction is as effective as engaging in mental imagery learning with an animated multimedia message to acquire stable transfer knowledge. The results of the present thesis provide first empirical data on long-term effects of drawing and show that drawing can be an effective strategy to acquire stable deeper understanding of a learning content. This finding is in line with research investigating long-term effects in multimedia learning indicating that learning with verbal and pictorial representations leads to stable knowledge gains (e.g., Schweppe, Eitel, & Rummer, 2015). The present results are in line with this research in that learning with provided and self-generated visualization lead to a stable benefit over learning without visual representations (Study 1), and learning with visualizations resulted in the acquisition of stable comprehension over time (Study 2). However, more empirical data is needed to broaden the body of evidence of how drawing affects learning over time. Moreover, the results of Study 1 and Study 2 were in line with the assumption that forgetting would occur for recognition performance over time. This finding indicates that propositional text features are stored only superficially in memory.

Fourth, within the second research question it was investigated whether benefits of drawing depend on the type of instructional support. The results of Study 3 showed that neither baseline nor extensive instructional support increased learning beyond an effect of unsupported drawing. The results indicate that support does not always constrain and structure the drawing process in a beneficial way. It seems that learners' reasoning about the elements they want to incorporate in their drawings is not positively affected by providing them with less relevant elements of the drawings. During drawing learners select and organize the most important information stated in the text (van Meter & Firetto, 2013). These processes require learners already to differentiate between relevant and less relevant elements. Thus, learners might ignore less relevant information early on during constructing drawings and do not reason about their visual appearance or spatial relations. So it seems plausible to argue that provided less relevant objects interfere with learners' own mental representations built during initial reading. This finding is in line with previous research on multimedia learning showing that learning can be hindered if learners build mental images during initial reading which interfere with processing provided illustrations (De Beni, Pazzaglia, Gyselinck, & Meneghetti, 2005; Pazzaglia & Cornoldi, 1999). When learners do not reason about these less relevant elements at all constructing a drawing that matches these elements might be rather demanding. Thus, additional reasoning might have been necessary in order to overcome the discrepancies between the provided visualizations and learners' own conceptions, thereby leaving less cognitive resources to engage in meaningful learning. Moreover, freeing learners from the need to reason about the visual appearance of any of the drawings' elements was not beneficial; a finding that is not in line with previous research (Lesgold et al., 1975, 1977; Schwamborn et al., 2011). Meaningful elements as well as distractors were presented from which learners could choose in order to generate drawings by spatially arranging these elements. Although it was assumed that reasoning about the relevance of elements should facilitate selecting processes that are important for learning with learner-generated as well as provided visual representations (Mayer, 2014a; van Meter & Firetto, 2013), the results did not reflect this assumption. This finding is in line with previous research showing that drawing in terms of selecting and arranging elements was only beneficial for prose learning in first-graders when distractors were omitted and learners were provided with the correct elements to construct drawings (Lesgold et al., 1975). The present findings indicate that selecting processes were not facilitated by adding distractor items. It rather seems that learners were not able to distinguish between useful and not useful elements reflected by impaired drawing quality. It might be that selecting processes were triggered that did not match the selecting processes described within the CTML or CMDC. In a first step, learners have to select the most important information, thereby discriminating between what is relevant and what is less relevant for learning (Mayer 2009, 2014a, van Meter & Firetto, 2013). This is fundamentally different from discriminating between elements and distractors that both illustrate important parts of the drawings but differ in whether they are useful in arranging a correct drawing. The resulting additional reasoning might consume cognitive resources which cannot be invested in meaningful learning. Thus, distractors could be useful that illustrate less relevant elements instead of incorrect relevant elements of the drawings in order to facilitate selecting processes in line with the CTML and CMDC. Further research is needed to investigate the characteristics of helpful drawing aids and the circumstances under which instructional support is beneficial for learner-generated drawing.

Moreover, the results of Study 3 may contribute to the discussion of what processes should be mandatory in order for a learning activity to be considered drawing. On the one hand, one can argue that drawing in terms of learner-generated drawing includes all strategies in which learners engage in constructing an external visuo-spatial representation irrespective of learners' engagement in actual drawing. This includes copying relevant elements from a legend containing all relevant elements of a drawing (Schmeck et al., 2014), or selecting and combining provided elements to construct drawings (Lesgold et al., 1975, 1977; Schwamborn et al. 2011). On the other hand it can be argued that only those strategies should be considered drawing in which students engage in actual drawing. This includes drawing instructions in which learners engage in free-hand drawing, for example when learners construct a drawing from scratch (e.g., Scheiter et al., 2017; Leopold & Leutner, 2012) or when they complete provided pre-drawn backgrounds (Study 3). It seems plausible to claim that an instruction, which contains provided visual representations of relevant elements of the drawings involve other cognitive processes than self-generating drawings from scratch, for example applying prior knowledge is not that relevant when the visuo-spatial appearance of relevant elements is derived during mental model construction.

To sum up, the present research substantiates previous findings that the type of posttest and the quality of constructed drawings during learning are important boundary conditions of learner-generated drawing. Benefits of drawing are more likely to be revealed on assessments of higher-order knowledge that match the characteristics of the mental model constructed during drawing. Moreover, drawing quality – as indicated by the number of main concepts stated in the text correctly incorporated into the drawings –is positively associated with test performance; also described as the prognostic drawing effect (Schwamborn et al., 2010). The prognostic drawing effect might also be affected by the type of posttest. However, there is little evidence in favor of this assumption resulting from the present research which is likely due to the way how transfer performance was assessed in Study 3. Additionally, the present results indicate that test delay and instructional support do not affect benefits of drawing under all circumstances. More research is needed to determine, clarify, and specify boundary conditions for learner-generated drawing.

Furthermore, boundary conditions of multimedia learning may also influence benefits of drawing. This assumption seems reasonable because visualization and externalization were found to be the main contributing factors to benefits of drawing in the present research. In line with research on multimedia learning, effects of learner-generated drawing could also depend on the degree of spatial information conveyed in the text (Schmidt-Weigand & Scheiter, 2011). In their experiment, Schmidt-Weigand and Scheiter (2011) found a multimedia effect only for low spatial but not for high spatial text. This finding is in line with the redundancy effect (Kalyuga & Sweller, 2014; Sweller, 2005b) suggesting that learning will be hindered when identical information is conveyed through multiple representations (i.e., through text and visualizations). Thus, benefits of drawing may also depend on how much spatial information is conveyed through the text. On the one hand, a high spatial text (i.e., a text containing a high degree of spatial information) may bare the risk of a redundancy effect in that the text and the learner-generated drawing contain identical spatial information. On the other hand, a text conveying too little spatial information may insufficiently provide learners with information they need in order to be able to construct drawings that correctly incorporate the spatial relations of important text elements. Moreover, benefits of drawing may depend on the relative distance between provided text and learner-generated drawings in line with the spatial contiguity principle (Mayer, 2001, 2009) suggesting that learning is fostered when corresponding text and visualizations are presented near rather than far from each other on the same page. Thus, it may be that learner-generated drawing is more beneficial when drawings are constructed on the same page on which the corresponding text is presented rather than when drawings are constructed on a separate page. Lastly, individual differences, for example, in terms of learners' representational competence may not only influence multimedia learning (Renkl & Scheiter, 2015) but may also affect whether learners successfully engage in a drawing strategy instruction. Representational competence includes a number of skills for constructing, interpreting, transforming and coordinating external representations, as well as domain-specific representational conventions (Kozma & Russell, 1997; Renkl & Scheiter, 2015). For example, learners who lack representational competence selected visualizations that were not well suited for a certain task or failed to coordinate the use of multiple representations (Stieff, Hegarty, & Deslongchamps, 2011). If learners lack domain-specific conventions or skills for constructing representations, they might generate poor quality drawings in terms of visually and spatially correct incorporated main concepts. To conclude, the present finding that benefits of drawing mainly stem from the availability of an external visual representation during learning suggests that some boundary conditions of multimedia learning may also affect benefits of learner-generated drawing. Further research is needed to determine whether this assumption holds true.

#### 4.2.4 How does drawing influence perceived difficulty of the instructional material?

The third research question examined how drawing affects learners' perceived difficulty of learning. The results of Study 2 revealed that drawing increased perceived difficulty. This finding is in line with previous research (Leutner et al., 2009) supporting the assumption that drawing without additional instructional support rises perceived difficulty because of increased cognitive demands concerned with managing the mechanics and logistics of constructing drawings. Learners may have experienced the task as being too difficult and, in turn, stopped at investing further effort in meaningful learning processes (Schmeck et al., 2015). Instructional support was assumed to have a positive influence on drawing as it, among other things, reduces perceived difficulty (e.g., Schmeck et al., 2014). However, this assumption was not yet investigated in a more systematic way in that supported and unsupported drawing have not directly been compared to one another in previous research.

In Study 3 it was examined whether instructional support during drawing can decrease perceived difficulty compared to a free-hand drawing condition. The results showed that neither baseline nor extensive instructional support did reduce perceived difficulty but it rather increased perceived difficulty compared to learners who did not receive support during learning. The results suggest that learners receiving baseline support were still concerned with managing the logistics and mechanics of constructing drawings. Moreover, this finding also supports the assumption that learners first needed to overcome their own conceptions built during initial reading which interfered with the representations used in the pre-drawn backgrounds. Thus, additional reasoning is necessary to overcome the discrepancies between learners mental images and the provided visual representations which hinders learning (De Beni et al., 2005; Pazzaglia & Cornoldi, 1999). The results of Study 3 also indicate that differentiating between relevant elements and non-useful distractors results in higher cognitive demands for adult learners rather than facilitates selecting processes that are assumed to promote learning with external pictorial representations (Mayer, 2009, 2014a; van Meter & Firetto, 2013; van Meter & Garner, 2005). It is recommended to further investigate the influence of distractors for this type of support to gain more insight in the underlying processes that are involved.

Unexpectedly, drawing without support did not increase perceived difficulty as suggested by the results of Study 2 and previous research (Leutner et al., 2009). The lack of a text-only control condition in Study 2 makes it difficult to compare the perceived difficulty ratings of Study 2 to the ratings of Study 3. The finding of Study 3 that the no-support drawing condition reported similar perceived difficulty ratings than the control condition could depend on the learning material. Whereas Leutner et al. (2009) used a learning material that includes more abstract visual representations of molecules, the learning material used in the presented study mainly included processes predominantly involving objects that are visible to the human eye. It seems plausible to argue that learners could more heavily rely on prior knowledge regarding the visual appearance of these objects (e.g., the human body or a hand). Thus, learners were not that negatively affected by free-hand drawing as they were in the study of Leutner et al. (2009). Moreover, this explanation also aligns with the assumption that the representations used for instructional support interfered with the learners' own representations built during initial reading as they may have been heavily influenced by learners' prior knowledge.

## 4.3 Practical Implications

From a practical point of view, learner-generated drawing seems to be a powerful strategy to learn from expository text in order for learners to gain a deeper understanding of the content. Educational practitioners could easily apply this strategy in the classroom as no additional preparation of material is needed (e.g., designing visual representations or graphs accompanying the text) besides providing students with an expository text. However, teachers should consider at least three things before instructing their students to engage in learner-generated drawing. First, the quality of drawings constructed during learning predicts knowledge acquisition. Thus, it is recommended to teachers to invest a course unit on how to construct a high quality learner-generated drawing in order for students to benefit from this strategy. Second, because drawing promotes processing and integration of visuo-spatial information, learning materials of natural science domains like biology, chemistry, or physics seem to be most suitable for a drawing strategy instruction. Lastly, teachers need to consider whether the kind of knowledge they want to convey can be acquired through a drawing instruction. Moreover, teachers need to use a test to assess learning outcomes that (a) measures what has to be taught according to the educational standards, and that (b) is suitable to capture the kind of knowledge that is built during drawing. Because of the characteristics of the mental model and perceptual image – two important mental representations built during drawing – tests involving problem-solving tasks, comprehension or transfer questions, as well as drawing tasks are most eligible for measuring the kind of knowledge that is built during learner-generated drawing.

Although drawing seems to be as effective as learning with provided visual representations, it can be recommended to use learner-generated drawing in the classroom instead of provided illustrations. In preparing a lesson, the teacher must decide which materials are well designed and which are not. It can be difficult to identify a well-designed multimedia message. There are a bunch of design principles for a multimedia message to be most effective for learning (Mayer, 2005, 2014b). However, schoolbooks often tend to include multimedia materials that are not in line with these assumptions (e.g., illustrations of a paragraph are presented on the next pages and not next to the paragraph, c.f. spatial contiguity principle; Mayer, 2001, 2009) and thus, are not well suited for learning. Therefore, it seems to be more efficient in terms of preparing a school lesson to use a comprehensible expository text and instruct students to self-generate accompanying illustrations. Moreover, other variables than learning outcome may be considered to determine whether teachers should use a drawing strategy instruction rather than a multimedia message. A drawing strategy instruction has been found to promote interest and motivation (Costantino, 1986; Fisher, 1976; Johnson, 1988; McConnell, 1993; Moore & Caldwell, 1993) and is assumed to promote learners' cognitive engagement (cf. Chi, 2009; Chi & Wylie, 2014) as well as integration of verbal and pictorial information (van Meter & Garner, 2005). Because learners tend to ignore provided pictures within a multimedia message (Hegarty & Just, 1993) combining a multimedia message with drawing instructions may be interesting. In this way, constructing a drawing should ensure that the pictorial information is processed thoroughly enabling learners to benefit from learning with external visual representations.

## 4.4 Strengths of the Present Dissertation

The present thesis is associated with multiple strengths concerning the quality of the conducted studies. First, learner-generated drawing was systematically compared to a series of learning strategies to investigate the influence of contributing factors on drawing. In an attempt to disentangle these effects, the focus was on only three contributing factors. In Study 1 and Study 2 one of these factors was hold constant. In Study 2, all learning strategies involved externalization while the factors generation and visualization were varied. In Study 2, visualization was held constant between conditions, while the factors generation and externalization were varied.

Another strength of the present studies was that learning outcome was measured with three different types of tests. The studies not only focused on assessments of higher-order knowledge, which are assumed to be affected by drawing based on recent theories and research findings, but also on assessments of lower-order knowledge that are assumed to be not affected by drawing (for overviews see Leutner & Schmeck 2014, van Meter & Firetto, 2013, van Meter & Garner, 2005). Lower-order assessments of recognition or free recall are widely used within educational research and practice but previous research and the present findings suggest that these assessments are not well suited to reveal benefits of drawing. It seems that benefits of drawing depend on a number of boundary conditions that should be further investigated in future studies. Moreover, the present research included not only one, but two assessments of higher-order knowledge, namely a transfer knowledge test and a drawing test to assess visuo-spatial knowledge.

A new learning material about biomechanics in human swimming behavior was constructed for the present thesis that was designed to be challenging for learners in order to obtain no ceiling effects resulting from university students being the subjects of the studies. The results of the studies can contribute to research on learner-generated drawing as they expand benefits of drawing to another content domain.

Lastly, the importance of support was investigated in a systematic way comparing a free-hand drawing group that received no support with two groups receiving additional instructional support during drawing and a non-drawing control condition. The types of support used in Study 3 built up on drawing support that was previously tested in empirical studies (Lesgold et al., 1977, 1975; Schmeck et al., 2014; Schwamborn et al., 2011) but was not yet compared to a drawing condition that received no instructional support or to each other.

## 4.5 Limitations

Despite the strengths of the present dissertation, there are also some limitations that need to be considered. First, no text-only condition was included in Study 2 that could have served as a control condition. The lack of this control condition may explain why there were no differences between conditions with regard to transfer performance. However, one cannot tell whether engaging in drawing, observing, or mental imagery was beneficial for learning compared to reading only.

Another possible limitation of the present thesis relates to the fact that no instructional support was used in the drawing conditions of Study 1 and Study 2. Previous research recommends that a drawing strategy instruction should contain some kind of support (van Meter & Firetto, 2013, van Meter & Garner, 2005); even some studies indicate that drawing is only effective if some kind of additional drawing aid is provided during learning. (van Meter, 2001; van Meter et al., 2006). Despite these previous research findings it was decided not to include additional drawing aids in Study 1 and Study 2 in order to investigate effects of learner-generated drawing without the influence of drawing aids. On the other hand, drawing support did not further increase learning beyond a general effect of free-hand drawing in Study 3. It is yet to determine which types of support are beneficial for drawing, and what moderating factors may influence the beneficial effect of support.

Moreover, the learning material that was used in the three empirical studies was rather long and complex. The text about biomechanics in human swimming behavior consisted of nine pages and contained multiple important concepts in the different paragraphs. This was in contrast to other learning materials used in previously conducted research on drawing which consisted of only two or three pages (e.g., van Meter, 2001), and of a fewer number of important concepts that were intertwined at the end of the text into a bigger picture (e.g., learning unit about the greenhouse effect; Scheiter et al., 2017).

Lastly, the samples that were used in all three studies limit the generalizability of the present findings. Whereas the present results may hold true for learners on a university level, it remains an open question whether the findings can be replicated with younger learners. Thus, it would be interesting to investigate students to derive implications for learning in the classroom.

### 4.6 Future Directions

Based on the present findings and the theoretical background provided in Chapter 1, implications for future directions in research on learner-generated drawing will be outlined in the following paragraph. First, to generalize the present findings it is recommended to replicate the studies taking the outlined limitations into account. Moreover, using other age groups and learning materials covering different content domains may be desirable in order to further substantiate the body of research on learner-generated drawing. Boundary conditions of multimedia learning, for example the redundancy principle (Kalyuga & Sweller, 2014; Sweller, 2005b), the spatial contiguity principle (Mayer, 2001, 2009), or the role of individual differences, may also affect learner-generated drawing and should be investigated in future research.

Additionally, motivational factors could be a moderating factor for benefits of drawing in line with assumptions of motivational factors influencing multimedia learning (Moreno & Mayer, 2007). One the one hand, the relevance of the learning topic could influence benefits of drawing. Learners might invest more effort in learning if the to-be-learned content is of personal relevance. This could be addressed in future studies investigating learner-generated drawing in a more applied context, for example with students in the classroom learning about a content that is part of the curriculum. On the other hand, learner’s goal orientation might also have an impact on benefits of drawing in line with assumptions involving the role of goal orientation as a motivational factor affecting self-regulated learning (Pintrich, 2000).

The CMDC also incorporates metacognitive processes derived from principles of self-regulated learning (Winne & Hadwin, 1998). Metacognition is assumed to be involved in learner-generated drawing, for example, when learners monitor their drawing activities and decide whether they can terminate their drawing activity or whether they need to go back to the perceptual image or mental model in order to revise their drawing (van Meter & Firetto, 2013). However, there is a lack of studies investigating metacognitive processes during drawing. Thus, future studies could focus on how drawing effects metacognition in terms of monitoring (i.e., the ability to assess one's state of learning compared to one's learning objectives), and control (i.e., the decision to terminate the learning activity or to restudy what has not yet been understood; Bjork, Dunlosky, & Kornell, 2013).

Lastly, another future direction of research on learner-generated drawing could head into the direction of using drawing for collaborative learning. Based on the assumptions of the ICAP framework (Chi, 2009; Chi & Wylie, 2014), learning through drawing could increase learners' cognitive engagement if learners' engage in an interactive discussion with a partner about their previously constructed drawings of the learning content. In line with the ICAP hypothesis (Chi & Wylie, 2014) higher cognitive engagement through discussing the drawings should foster learning.

## 4.7 Conclusions

The presented thesis aimed at investigating which factors mainly contribute to benefits of learner-generated drawings (van Meter & Firetto, 2013) on learning outcomes, what boundary conditions affects these benefits (type of posttest, drawing quality, test delay, instructional support), and how a drawing strategy instruction influences perceived difficulty of the learning content.

The results of Study 1 and Study 2 indicate that benefits of drawing mainly stem from processing an externalized visual representation of the learning content. The process of generation seems to have no additional benefit. The results of all three studies replicated the finding from previous research that benefits of drawing are more likely to be revealed on assessments of higher-order knowledge than on assessments of lower-order knowledge (Leutner & Schmeck, 2014; van Meter & Firetto, 2013; van Meter & Garner, 2005). Moreover, a test that requires learners to apply visuo-spatial knowledge to new contexts was used that extends the body of suitable higher-order assessments. The results of Study 2 also suggest that the size of the prognostic drawing effect (Schwamborn et al., 2010) is similarly affected by the type of posttest, that is, the positive correlation between drawing quality and learning outcomes being higher for assessments of higher-order knowledge than for assessments of lower-order knowledge. The lack of replicating this finding in Study 3 might be due to the way transfer performance was assessed. The multiple-choice questions used in the transfer test of Study 3 might have underestimated learners' knowledge of the content because learners were not able to express their understanding in more detail. However, there is little evidence that drawing constitutes a desirable difficulty (Bjork & Bjork, 2011; Bjork, 1994) affecting long-term effects of drawings. Instructional support did not further improve learning beyond a beneficial effect of free-hand drawing. Thus it seems that drawing aids are not helpful under all circumstances and may depend on (a) their design, and (b) the learning material. In line with previous research (Leutner et al., 2009), free-hand drawing increased perceived difficulty in Study 2 – a finding that could not be replicated in Study 3. Instructional support did not decrease – as would have been expected according to previous studies (Schmeck et al., 2014, Schwamborn et al., 2011) – but rather than increased perceived difficulty. These findings suggest that the relation between drawing and perceived difficulty seems to be more complex and should be addressed in future research.

To conclude, visualization and externalization seem to be the main contributing factors to benefits of learner-generated drawing. Thus, recent theoretical frameworks of drawing (e.g., the CMDC; van Meter & Firetto, 2013) may overemphasize the role of generation and the results of previous research comparing a drawing group to a non-drawing control group might be interpreted differently. Further research is needed to get more insight on long-term effects of drawing as well as the design of instructional support and its boundary conditions to be effective to support learners in drawing construction.

# 5 Summary

Learning with visual representations has been the focus of research for the past decades. A robust finding is the multimedia effect (Butcher, 2014), that is, learning with a combination of text and pictures is more beneficial than learning with text alone. Three cognitive processes are assumed to be important for multimedia learning: selecting and organizing information from text and pictures, as well as integrating this information in order to build a coherent mental model of the learning content (Mayer, 2009, 2014a). Although integration is crucial for effective multimedia learning, learners show only few attempts to integrate (Hegarty & Just, 1993; Mason et al., 2013, 2015; Scheiter & Eitel, 2015). One way to foster integration of verbal and pictorial information is to instruct learners to self-generate the visual representations. This strategy is described as learner-generated drawing (van Meter & Firetto, 2013, van Meter & Garner, 2013).

Learner-generated drawing requires learners to construct external pictorial representations that include the key concepts and their relations while learning from verbal instruction (Leutner & Schmeck, 2014). Drawing has been shown to promote learning in terms of higher-order knowledge (for overviews see Ainsworth et al., 2011; van Meter & Firetto, 2013; van Meter & Garner, 2005). Benefits of drawing seem to depend on the availability and type of support during drawing construction (e.g., van Meter, 2001, van Meter et al., 2006, Schwamborn et al., 2010, Schmeck et al., 2014). Moreover, the quality of drawings constructed during drawing (i.e., the number of key concepts correctly incorporated into the drawings) is positively associated with learning outcomes; a finding that can be described as the prognostic drawing effect (Schwamborn et al., 2010). The main goal of the present thesis was to investigate which of three factors – generation, visualization, and externalization – mainly contributes to benefits of drawing. Second, it was examined how several boundary conditions would affect benefits of drawing – namely, the type of posttest, the quality of drawings constructed during learning, test delay as well as type of instructional support. Third, it was of interest how a drawing strategy instruction influences learners' perceived difficulty of learning.

The first study investigated the influence of generation and visualization. Drawing was compared with a multimedia condition, a summary condition, and a text-only condition. After learning, learning outcomes were assessed with an immediate and delayed posttest, whereby the time of assessment was manipulated within subjects. The results indicate that visualization is the factor contributing most to benefits of drawing. Because the results regarding time of testing likely were confounded with benefits of retrieval practice (Roediger & Karpicke, 2006), time of testing was included as a between-subjects factor in the second study.

The second study examined the influence of generation and externalization on benefits of drawing. To this end, drawing was compared with an observation condition and a mental imagery condition. The results of the first study were replicated regarding the influence of an external pictorial representation being the main contributing factor to benefits of drawing. Moreover, it seems that drawing does not constitute a desirable difficulty (Bjork & Bjork, 2011; Bjork, 1994). As expected, the results showed that the size of the prognostic drawing effects depends on the type of posttest in that it was larger for assessments of higher-order knowledge than for assessments of lower-order knowledge. Additionally, the results indicate that free-hand drawing left learners with less cognitive resources available to engage in meaningful learning.

The third study investigated which type of instructional support is most effective in order to help learners benefit more from constructing drawings. To this end, a no-support drawing condition was contrasted to a low-support condition, a high-support condition, and a text-only control condition. The findings indicate that neither reducing the requirement to reason about irrelevant elements of the drawings alone, nor reducing the need to reason about the visual appearance of any element seemed to influence learning beyond an unsupported drawing effect. Moreover, instructional support increased rather than decreased cognitive demands associated with managing the drawing process. Drawing quality was positively associated with learning outcomes; however, the prognostic drawing effect was not larger for assessments of higher-order knowledge than for assessments of lower-order knowledge.

In conclusion, benefits of drawing seem to stem mainly from externalizing a visualization that drawing requires, rather than the actual generation of the drawing. Accordingly, recent theoretical frameworks of drawing (e.g., the CMDC; van Meter & Firetto, 2013) may overemphasize the role of generation. Thus, the results of previous studies comparing a drawing group to a non-drawing control group might be interpreted differently. Further research is needed to get more insight on boundary conditions of drawing including long-term effects, the influence of the type of posttest on the prognostic drawing effect, and the design of beneficial instructional support, as well as the influence of perceived difficulty for learner-generated drawing. Moreover, boundary conditions of multimedia learning could also affect benefits of learner-generated drawing and should be considered in future studies.

# 6 Zusammenfassung

Der Einsatz von visuellen Repräsentationen beim Lernen wird seit einigen Jahrzehnten wissenschaftlich untersucht. Dabei stellt der so genannte Multimedia Effekt einen stabilen Befund dar (Butcher, 2014): Das Lernen mit einer Kombination aus Text und Bildern ist vorteilhafter als das Lernen mit reinem Text. Für multimediales Lernen werden drei zentrale kognitive Prozesse unterschieden: Selektion und Organisation wichtiger Informationen aus Text und Bildern, sowie Integration dieser Informationen um ein kohärentes mentales Modell des Lerninhalts zu bilden (Mayer, 2009, 2014a). Obwohl vor allem der Integrations-Prozess eine besonders wichtige Rolle bei effektivem multimedialem Lernen spielt, zeigen Lernende nur wenige Anläufe die Informationen aus den verschiedenen Quellen zu integrieren (Hegarty & Just, 1993; Mason et al., 2013, 2015; Scheiter & Eitel, 2015). Eine Möglichkeit zur Förderung der Integration von verbalen und piktorialen Informationen ist Lernende zu instruieren, die visuellen Repräsentationen selbst zu generieren. Diese Strategie wird auch als Anfertigen selbst generierter Zeichnungen beschrieben (van Meter & Firetto, 2013, van Meter & Garner, 2013).

Das Anfertigen selbst generierter Zeichnungen erfordert, dass Lernende externale piktoriale Repräsentationen eines Lerninhalts anfertigen, welche die wichtigsten Konzepte und deren räumliche Beziehungen zueinander abbilden, die in einem Text beschrieben werden. (Leutner & Schmeck, 2014). Zeichnen kann den Erwerb von Wissen höherer Ordnung (z.B. Transferwissen) fördern (für einen Überblick siehe Ainsworth et al., 2011; van Meter & Firetto, 2013; van Meter & Garner, 2005). Die Vorteile der Zeichenstrategie scheinen von der Verfügbarkeit und der Art der instruktionalen Unterstützung während des Zeichnens abzuhängen (van Meter, 2001, van Meter et al., 2006, Schwamborn et al., 2010, Schmeck et al., 2014). Darüber hinaus bestimmt die Qualität der während des Lernens angefertigten Zeichnungen (d.h. die Anzahl der wichtigen Konzepte, die korrekt in die

Zeichnungen überführt wurden) den anschließenden Lernerfolg. Dieser Befund wird auch prognostischer Effekt des Zeichnens genannt (Schwamborn et al., 2010). Das Hauptziel der vorliegenden Dissertation war die Untersuchung, welcher von drei Faktoren – Generierung, Visualisierung und Externalisierung – am stärksten zu förderlichen Effekten des Zeichnens beiträgt. Außerdem wurde erforscht, wie verschiedene Randbedingungen – und zwar die Art des Posttests, die Zeichnungsqualität, verzögertes Testen, sowie die Art der instruktionalen Unterstützung – Vorteile der Zeichenstrategie beeinflussen. Zuletzt war von Interesse, wie eine Zeichnen-Instruktion die wahrgenommene Schwierigkeit der Lernenden beeinflusst.

Die erste Studie untersuchte den Einfluss von Generierung und Visualisierung. Eine Zeichnen-Bedingung wurde mit einer Multimedia-Bedingung, einer Zusammenfassung-Bedingung und einer reinen Text-Bedingung verglichen. Nach der Lernphase wurde das Lernergebnis mit einem unmittelbaren und einem verzögerten Posttest abgefragt, wobei der Testzeitpunkt innerhalb der Versuchspersonen manipuliert wurde. Die Ergebnisse deuten darauf hin, dass Visualisierung der Faktor ist, der am stärksten zu den förderlichen Effekten des Zeichnens beiträgt. Da die Ergebnisse hinsichtlich des Testzeitpunkts durch Vorteile des wiederholten Abrufens (Roediger & Karpicke, 2006) konfundiert sein könnten, wurde der Testzeitpunkt in der zweiten Studie zwischen den Versuchspersonen manipuliert.

Die zweite Studie fokussierte auf den Einfluss von Generierung und Externalisierung auf positive Effekte des Zeichnens. Dazu wurde eine Zeichnen-Bedingung mit einer Beobachtungs-Bedingung und einer Vorstellungs-Bedingung verglichen. Die Ergebnisse der ersten Studie wurden dahingehend repliziert, dass das Vorhandensein einer externalen piktorialen Repräsentation der entscheidende Einflussfaktor für positive Effekte des Zeichnens zu sein scheint. Darüber hinaus deutet wenig darauf hin, dass das Anfertigen von Zeichnungen eine wünschenswerte Erschwernis während des Lernens darstellt (Bjork & Bjork, 2011; Bjork, 1994). Wie erwartet zeigten die Ergebnisse, dass die Größe des prognostischen Zeichnen-Effekts von der Art des Posttests abhängt: der prognostische Effekt des Zeichnens war größer für Maße, die Wissen höherer Ordnung erfassen (z.B. Transferleistung) als für Maße, die Wissen niedrigerer Ordnung erfassen (z.B. Wiedererkennungsleistung). Zudem deuten die Ergebnisse darauf hin, dass Lernenden beim Anfertigen freihändiger Zeichnungen weniger kognitive Ressourcen zur Verfügung stehen um bedeutungsvollen Lernprozessen nachzugehen.

Die dritte Studie beschäftigte sich damit, welche Arten instruktionaler Unterstützung am wirkungsvollsten sind um Lernenden zu helfen, stärker vom Anfertigen von Zeichnungen zu profitieren. Zu diesem Zweck wurde eine „keine Unterstützung"-Zeichnenbedingung mit einer „niedrige Unterstützung"-Zeichnenbedingung, einer „hohe Unterstützung"-Zeichnenbedingung und einer reinen Text-Bedingung verglichen. Die Befunde deuten darauf hin, dass weder das Reduzieren der Notwendigkeit über weniger relevante Elemente der Zeichnungen zu schlussfolgern, noch die Reduzierung der Notwendigkeit über das Erscheinungsbild jeglicher Elemente der Zeichnungen zu schlussfolgern das Lernergebnis so beeinflussen, dass dies über einen positiven Effekt des nicht-unterstützten Zeichnens hinaus gefördert wird. Außerdem erhöhte instruktionale Unterstützung die kognitiven Anforderungen, die mit dem Steuern des Zeichenprozesses assoziiert sind, anstatt diese zu reduzieren. Die Zeichnungsqualität war positiv mit dem Lernergebnis verbunden, jedoch war der prognostische Zeichnen-Effekt für Maße, die Wissen höherer Ordnung erfassen nicht größer als für Maße, die Wissen niedrigerer Ordnung erfassen.

Die Gesamtheit der Ergebnisse legt nahe, dass die Vorteile einer Zeichenstrategie hauptsächlich aus der Externalisierung einer visuellen Repräsentation resultieren, anstatt aus der tatsächlichen Generierung einer Zeichnung. Dementsprechend könnten jüngste theoretische Modelle des Zeichnens (z.B. das CMDC; van Meter & Firetto, 2013) die Bedeutung von Generierung überbetonen. Somit sollten die Ergebnisse früherer Studien, die eine Zeichnen-Bedingung mit einer reinen Text-Bedingung vergleichen, möglicherweise auf eine andere Art und Weise interpretiert werden. Weitere Forschung ist nötig um mehr Einblick in den Zusammenhang zwischen wahrgenommener Schwierigkeit des Lernens und positiven Effekten des Zeichnens zu erhalten. Zudem sollten die Rahmenbedingungen weiter untersucht werden, welche die positiven Effekte des Zeichnens beeinflussen. Zu diesen gehören Langzeit-Effekte, der Einfluss der Art des Posttests auf den prognostischen Zeichnen-Effekt und die Gestaltung wirksamer instruktionaler Unterstützungsmaßnahmen. Darüber hinaus könnten die Rahmenbedingungen für erfolgreiches multimediales Lernen auch beim Anfertigen selbst generierter Zeichnungen eine Rolle spielen.

# 7 References

Ainsworth, S. (1999). The functions of multiple representations. *Computers & Education*, *33*, 131–152. http://doi.org/10.1016/s0360-1315(99)00029-9

Ainsworth, S., Prain, V., & Tytler, R. (2011). Drawing to learn in science. *Science*, *333*, 1096–1097. http://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. http://doi.org/10.1037/0022-0663.73.3.358

Alesandrini, K. L. (1984). Pictures and adult learning. *Instructional Science*, *13*, 63–77. http://doi.org/10.1007/BF00051841

Arcavi, A. (2003). The role of visual representations in the learning of mathematics. *Educational Studies in Mathematics*, *52*, 215–241. http://doi.org/10.1023/A:1024312321077

Atkinson, R. C., & Shiffrin, R. M. (1968). Human memory: A proposed system and its control processes. In K. W. Spence & J. T. Spence (Eds.), *The psychology of learning and motivation* (Volume 2, pp. 89–195). New York: Academic Press. http://doi.org/http://doi.org/10.1016/S0079-7421(08)60422-3

Baddeley, A. D. (1992). Working memory. *Science*, *255*, 556–559. http://doi.org/10.1126/science.1736359

Baddeley, A. D. (1999). *Human memory*. Boston: Allyn & Bacon.

Baddeley, A. D., Eysenck, M. W., & Anderson, M. C. (2015). *Memory* (Second edition). East Sussex: Psychology Press.

Bean, T. W., & Steenwyk, F. L. (1984). The effect of three forms of summarization on sixth graders' summary writing and comprehension. *Journal of Reading Behavior*, *16*, 298–306.

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 US: Worth Publishers.

Bjork, R. A. (1994). Memory and metamemory considerations in the training of human beings. In J. Metcalfe & A. P. Shimamura (Eds.), *Metacognition: Knowing about knowing*. (pp. 185–205). Cambridge, MA: MIT Press.

Bjork, R. A., Dunlosky, J., & Kornell, N. (2013). Self-regulated learning: Beliefs, techniques, and illusions. *Annual Review of Psychology*, *64*, 417–444. http://doi.org/10.1146/annurev-psych-113011-143823

Bodemer, D., Ploetzner, R., Feuerlein, I., & Spada, H. (2004). The active integration of information during learning with dynamic and interactive visualisations. *Learning and Instruction*, *14*, 325–341. http://doi.org/10.1016/j.learninstruc.2004.06.006

Britton, L. A., & Wandersee, J. H. (1997). Cutting up text to make moveable, magnetic diagrams: A way of teaching and assessing biological processes. *The American Biology Teacher*, *59*, 288–291.

Butcher, K. R. (2006). Learning from text with diagrams: Promoting mental model development and inference generation. *Journal of Educational Psychology*, *98*, 182–197. http://doi.org/10.1037/0022-0663.98.1.182

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

Caldwell, H., & Moore, B. H. (1991). The art of writing: Drawing as preparation for narrative writing in the primary grades. *Studies in Art Education*, *32,* 207–219.

Carney, R. N., & Levin, J. R. (2002). Pictorial illustrations still improve students' learning from text. *Educational Psychology Review*, *14*, 5–26. http://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. http://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. http://doi.org/10.1080/00461520.2014.965823

Clark, J. M., & Paivio, A. (1991). Dual coding theory and education. *Educational Psychology Review*, *3*, 149–210. http://doi.org/10.1007/BF01320076

Cook, M., Wiebe, E. N., & Carter, G. (2008). The influence of prior knowledge on viewing and interpreting graphics with macroscopic and molecular representations. *Science Education*, *92*, 848–867. http://doi.org/10.1002/sce.20262

Costantino, T. M. (1986). Drawing: Homework for remedial readers. *The Reading Teacher*, *39*, 737–739.

Cox, R. (1999). Representation construction, externalised, cognition and individual differences. *Learning and Instruction*, *9*, 343–363. http://doi.org/10.1016/S0959-4752(98)00051-6

Cuevas, H. M., Fiore, S. M., & Oser, R. L. (2002). Scaffolding cognitive and metacognitive processes in low verbal ability learners: Use of diagrams in computer-based training environments. *Instructional Science*, *30*, 433–464. http://doi.org/10.1023/A:1020516301541

De Beni, R., Pazzaglia, F., Gyselinck, V., & Meneghetti, C. (2005). Visuospatial working memory and mental representation of spatial descriptions. *European Journal of Cognitive Psychology*, *17*, 77–95. http://doi.org/10.1080/09541440340000529 de Bock, D., Verschaffel, L., & Janssens, D. (1998). The predominance of the linear model in secondary school students' solutions of word problems involving length and area of similar plane figures. *Educational Studies in Mathematics*, *35*, 65–83. http://doi.org/10.1023/A:1003151011999

de Bock, D., Verschaffel, L., Janssens, D., van Dooren, W., & Claes, K. (2003). Do realistic contexts and graphical representations always have a beneficial impact on student's performance? Negative evidence from a study on modelling non-linear geometry problems. *Learning and Instruction*, *13*, 441–463. http://doi.org/10.1016/S0959-4752(02)00040-3

Dempsey, B. C., & Betz, B. J. (2001). A Scientific Tool for Learning. *The American Biology Teacher*, *63*, 271–281.

Dietz, S. (1976). Monsters!? *Teacher*, *93*, 64–65.

Dunlosky, J., Rawson, K. A., Marsh, E. J., Nathan, M. J., & Willingham, D. T. (2013). Improving students' learning with effective learning techniques: Promising directions from cognitive and educational psychology. *Psychological Science in the Public Interest*, *14*, 4–58. http://doi.org/10.1177/1529100612453266

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

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

Eitel, A., Scheiter, K., Schüler, A., Nyström, M., & Holmqvist, K. (2013). How a picture facilitates the process of learning from text: Evidence for scaffolding. *Learning and Instruction*, *28*, 48–63. http://doi.org/10.1016/j.learninstruc.2013.05.002

Ekstrom, R. B., French, J. W., & Harman, H. H. (1976). *Manual for kit of factor-referenced cognitive tests*. Princeton, NJ: Educational Testing Service.

Fisher, L. J. (1976). Language arts: Pictures tell the tale. *Teacher, 93*, 64–73.

Foos, P. W., Mora, J. J., & Tkacz, S. (1994). Student study techniques and the generation effect. *Journal of Educational Psychology*, *86*, 567–576. http://doi.org/10.1037/0022-0663.86.4.567

Glenberg, A. M., & Langston, W. E. (1992). Comprehension of illustrated text: Pictures help to build mental models. *Journal of Memory and Language*, *31*, 129–151. http://doi.org/10.1016/0749-596X(92)90008-L

Gobert, J. D., & Clement, J. J. (1999). Effects of student-generated diagrams versus student-generated summaries on conceptual understanding of causal and dynamic knowledge in plate tectonics. *Journal of Research in Science Teaching*, *36*, 39–53. http://doi.org/10.1002/(SICI)1098-2736(199901)36:1<39::AID-TEA4>3.0.CO;2-I

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. http://doi.org/10.1016/0022-0965(89)90041-6

Grossen, B., & Carnine, D. (1990). Diagramming a logic strategy: Effects on difficult problem types and transfer. *Learning Disability Quarterly*, *13*, 168–182. http://doi.org/10.2307/1510699

Hall, V. C., Bailey, J., & Tillman, C. (1997). Can student-generated illustrations be worth ten thousand words? *Journal of Educational Psychology*, *89*, 677–681. http://doi.org/10.1037/0022-0663.89.4.677

Hannus, M., & Hyönä, J. (1999). Utilization of illustrations during learning of science textbook passages among low- and high-ability children. *Contemporary Educational*

*Psychology*, *24*, 95–123. http://doi.org/10.1006/ceps.1998.0987

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*, 92–102. http://doi.org/10.1037/0022-0663.89.1.92

Hegarty, M., & Just, M. A. (1993). Constructing mental models of machines from text and diagrams. *Journal of Memory and Language*, *32*, 717–742.

Hembree, R. (1992). Experiments and relational studies in problem solving : A meta-analysis. *Journal for Research in Mathematics Education*, *23*, 242–273.

Hubbard, R. S. (1987). Transferring images: Not just glued on the page. *Young Children*, *42*, 60–67.

Jairam, D., & Kiewra, K. A. (2010). Helping students soar to success on computers: An investigation of the SOAR study method for computer-based learning. *Journal of Educational Psychology*, *102*, 601–614. http://doi.org/http://dx.doi.org/10.1037/a0019137

Johnson, D. B. (1988). Show me what you mean: Student posters teach lengthy material. *Exercise Exchange*, *34*, 44–46.

Kalyuga, S., & Sweller, J. (2014). The redundancy principle in multimedia learning. In R. E. Mayer (Ed.), *The Cambridge handbook of multimedia learning* (2nd edition, pp. 247–262). Cambridge, NY: Cambridge University Press. http://doi.org/10.1017/CBO9781139547369.013

Karnowski, L. (1986). How young writers communicate. *Educational Leadership*, *44*, 58–60.

Kerr, R., & Booth, B. (1978). Specific and varied practice of motor skill. *Perceptual and Motor Skills*, *46*, 395–401.

Kozma, R. B., & Russell, J. (1997). Multimedia and understanding: Expert and novice responses to different representations of chemical phenomena. *Journal of Research in Science Teaching, 34*, 949–968. http://doi.org/10.1002/(SICI)1098-2736(199711)34:9<949::AID-TEA7>3.0.CO;2-U

Lansing, K. M. (1984). The effect of drawing on the development of mental representations: A continuing study. *Studies in Art Education*, *25*, 167–175.

Larkin, J. H., & Simon, H. A. (1987). Why a diagram is (sometimes) worth ten thousand words. *Cognitive Science, 11*, 65–99.

Leopold, C., & Leutner, D. (2012). Science text comprehension: Drawing, main idea selection, and summarizing as learning strategies. *Learning and Instruction*, *22*, 16–26. http://doi.org/10.1016/j.learninstruc.2011.05.005

Leopold, C., & Leutner, D. (2015). Improving students' science text comprehension through metacognitive self-regulation when applying learning strategies. *Metacognition and Learning*, *10*, 313–346. http://doi.org/10.1007/s11409-014-9130-2

Leopold, C., & Mayer, R. E. (2014). An imagination effect in learning from scientific text. *Journal of Educational Psychology*, *107*, 47–63. http://doi.org/10.1037/a0037142

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. http://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. http://doi.org/10.1080/10862967709547240.

Lesgold, A. M., Shimron, J., Levin, J. R., & Guttmann, J. (1975). Pictures and young children's learning from oral prose. *Journal of Educational Psychology*, *67*, 636–642. http://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. http://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 edition, pp. 433–448). Cambridge, NY: Cambridge University Press. http://doi.org/10.1017/CBO9781139547369.010

Liepmann, D., Beauducel, A., Brocke, B., & Amthauer, R. (2007). *Intelligenz-Struktur-Test 2000 R* (2nd, revised 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. http://doi.org/10.1016/j.cedpsych.2013.04.001

Mason, L., Tornatora, M. C., & Pluchino, P. (2013). Do fourth graders integrate text and picture in processing and learning from an illustrated science text? Evidence from eye-movement patterns. *Computers and Education, 60,* 95–109. http://doi.org/10.1016/j.compedu.2012.07.011

Mason, L., Tornatora, M. C., & Pluchino, P. (2015). Integrative processing of verbal and graphical information during re-reading predicts learning from illustrated text: An eye-movement study. *Reading and Writing, 28,* 851–872. http://doi.org/10.1007/s11145-015-9552-5

Mayer, R. E. (1989). Systematic thinking fostered by illustrations in scientific text. *Journal of Educational Psychology, 81*, 240–246. http://doi.org/10.1037/0022-0663.81.2.240

Mayer, R. E. (1997). Multimedia learning: Are we asking the right questions? *Educational Psychologist, 32,* 1–19. http://doi.org/10.1207/s15326985ep3201

Mayer, R. E. (2001). *Multimedia learning*. New York: Cambridge University Press.

Mayer, R. E. (2005). *The Cambridge handbook of multimedia learning*. New York, NJ:

Cambridge University Press.

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

Mayer, R. E. (2011). *Applying the science of learning*. Upper Saddle River, NZ: Pearson.

Mayer, R. E. (2014a). Cognitive theory of multimedia learning. In R. E. Mayer (Ed.), *The Cambridge handbook of multimedia learning* (2nd edition, pp. 43–71). Cambridge, NY: Cambridge University Press.

Mayer, R. E. (2014b). *The Cambridge handbook of multimedia learning* (2nd edition). Cambridge, NY: Cambridge University Press. http://doi.org/10.1017/CBO9781139547369

Mayer, R. E., & Anderson, R. B. (1991). Animations need narrations: An experimental test of a dual-coding hypothesis. *Journal of Educational Psychology*, *83*, 484–490. http://doi.org/10.1037//0022-0663.83.4.484

Mayer, R. E., & Anderson, R. B. (1992). The instructive animation: Helping students build connections between words and pictures in multimedia learning. *Journal of Educational Psychology*, *84*, 444–452. http://doi.org/10.1037/0022-0663.84.4.444

Mayer, R. E., & Gallini, J. K. (1990). When is an illustration worth ten thousand words? *Journal of Educational Psychology*, *82*, 715–726. http://doi.org/10.1037/0022-0663.82.4.715

Mayer, R. E., & Moreno, R. (2002). Animations as an aid to multimedia learning. *Educational Psychology Review*, *14*, 87–99.

McConnell, S. (1993). Talking drawings: A strategy for assisting learners. *Journal of Reading*, *36*, 260–269.

McNamara, T. P., Halpin, J. A., & Hardy, J. K. (1992). The representation and integration in memory of spatial and nonspatial information. *Memory & Cognition*, *20*, 519–532.

http://doi.org/10.3758/BF03199584

Miller, G. A. (1956). The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychological Review*, *63*, 81–97. http://doi.org/10.1037/h0043158

Moore, B. H., & Caldwell, H. (1993). Drama and drawing for narrative writing in primary grades. *Journal of Educational Research*, *87*, 100–110.

Moreno, R., & Mayer, R. E. (2007). Interactive multimodal learning environments. *Educational Psychology Review*, *19*, 309–326. http://doi.org/10.1007/s10648-007-9047-2

O'Keefe, P. A., Letourneau, S. M., Homer, B. D., Schwartz, R. N., & Plass, J. L. (2014). Learning from multiple representations: An examination of fixation patterns in a science simulation. *Computers in Human Behavior*, *35*, 234–242. http://doi.org/10.1016/j.chb.2014.02.040

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. http://doi.org/10.1037/0022-0663.84.4.429

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

Paivio, A. (1991). Dual coding theory: Retrospect and current status. *Canadian Journal of Psychology*, *45*, 255–287. http://doi.org/10.1037/h0084295

Paivio, A. (2006). *Mind and its evolution: A dual coding approach*. Mahawah, NJ: Lawrence Erlbaum.

Pazzaglia, F., & Cornoldi, C. (1999). The role of distinct components of visuo- spatial working memory in the processing of texts. *Memory*, *7*, 19–41. http://doi.org/10.1080/741943715

Pintrich, P. R. (2000). The role of goal-orientation in self-regulated learning. In M. Boekaerts, P. R. Pintrich, & M. Zeidner (Eds.), *The handbook of self-regulation* (pp. 451–502). San Diego, CA: Academic Press. http://doi.org/10.1016/B978-012109890-2/50043-3

Plass, J. L., Chun, D. M., Mayer, R. E., & Leutner, D. (1998). Supporting visual and verbal learning preferences in a second-language multimedia learning environment. *Journal of Educational Psychology*, *90*, 25–36. http://doi.org/10.1037/0022-0663.90.1.25

Rasco, R. W., Tennyson, R. D., & Boutwell, R. C. (1975). Imagery instructions and drawings in learning prose. *Journal of Educational Psychology*, *67*, 188–192. http://doi.org/10.1037/h0077014

Rellensmann, J., Schukajlow, S., & Leopold, C. (2016). Make a drawing. Effects of strategic knowledge, drawing accuracy, and type of drawing on students' mathematical modelling performance. *Educational Studies in Mathematics*, 1–26. http://doi.org/10.1007/s10649-016-9736-1

Renkl, A., & Scheiter, K. (2015). Studying visual displays: How to instructionally support learning. *Educational Psychology Review*. http://doi.org/10.1007/s10648-015-9340-4

Rich, R. Z., & Blake, S. (1994). Using pictures to assist in comprehension and recall. *Intervention in School and Clinic*, *29*, 271–275. http://doi.org/10.1177/105345129402900504

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

Roediger, H. L., & Karpicke, J. D. (2006). The power of testing memory: Basic research and implications for educational practice. *Perspectives on Psychological Science*, *1*, 181–210. http://doi.org/10.1111/j.1745-6916.2006.00012.x

Roediger, H. L., Putnam, A. L., & Smith, M. A. (2011). Ten benefits of testing and their applications to educational practice. In J. P. Mestre & B. H. Ross (Eds.), *The psychology of learning and motivation: Cognition in education.* (Vol. 55, pp. 1–36). San Diego, CA US: Elsevier Academic Press. http://doi.org/10.1016/B978-0-12-387691-1.00001-6

Rohrer, D., & Taylor, K. (2007). The shuffling of mathematics problems improves learning. *Instructional Science*, *35*, 481–498. http://doi.org/10.1007/s11251-007-9015-8

Scaife, M., & Rogers, Y. (1996). External cognition: How do graphical representations work? *International Journal of Human-Computer Studies*, *45*, 185–213. http://doi.org/10.1006/ijhc.1996.0048

Scheiter, K., & Eitel, A. (2015). Signals foster multimedia learning by supporting integration of highlighted text and diagram elements. *Learning and Instruction*, *36*, 11–26. http://doi.org/10.1016/j.learninstruc.2014.11.002

Scheiter, K., Schleinschok, K., & Ainsworth, S. (2017). Why sketching may aid learning from science texts: Contrasting sketching with written explanations. *Topics in Cognitive Science*. http://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. http://doi.org/10.1016/j.cedpsych.2014.07.003

Schmeck, A., Opfermann, M., van Gog, T., Paas, F., & Leutner, D. (2015). Measuring cognitive load with subjective rating scales during problem solving: Differences between immediate and delayed ratings. *Instructional Science*, *43*, 93–114. http://doi.org/10.1007/s11251-014-9328-3

Schmidt-Weigand, F., & Scheiter, K. (2011). The role of spatial descriptions in learning from multimedia. *Computers in Human Behavior*, *27*, 22–28. http://doi.org/10.1016/j.chb.2010.05.007

Schmitt, N. (1996). Uses and abuses of coefficient alpha. *Psychological Assessment*, *8*, 350–353.

Schneider, W., Schlagmüller, M., & Ennemoser, M. (2007). *Lesegeschwindigkeits- und -verständnistest für die Klassen 6-12*. (M. Hasselhorn, H. Marx, & W. Schneider, Eds.). Göttingen: Hogrefe.

Schnotz, W. (2002). Towards an integrated view of learning from text and visual displays. *Educational Psychology Review*, *14*, 101–120. http://doi.org/10.1023/A:1013136727916

Schnotz, W. (2014). Integrated model of text and picture comprehension. In R. E. Mayer (Ed.), *The Cambridge handbook of multimedia learning* (2nd edition, pp. 72–103). Cambridge, NY: Cambridge University Press.

Schnotz, W., & Bannert, M. (2003). Construction and interference in learning from multiple representation. *Learning and Instruction*, *13*, 141–156. http://doi.org/10.1016/S0959-47520200017-8

Schüler, A. (2017). Investigating gaze behavior during processing of inconsistent text-picture information: Evidence for text-picture integration. *Learning and Instruction*, *49*, 218–231. http://doi.org/10.1016/j.learninstruc.2017.03.001

Schüler, A., Arndt, J., & Scheiter, K. (2015). Processing multimedia material: Does integration of text and pictures result in a single or two interconnected mental representations? *Learning and Instruction*, *35*, 62–72. http://doi.org/10.1016/j.learninstruc.2014.09.005

Schwamborn, 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. http://doi.org/10.1037/a0019640

Schwamborn, 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. http://doi.org/10.1016/j.chb.2010.05.028

Schweppe, J., Eitel, A., & Rummer, R. (2015). The multimedia effect and its stability over time. *Learning and Instruction, 38*, 24–33. http://doi.org/10.1016/j.learninstruc.2015.03.001

Seufert, T. (2003). Supporting coherence formation in learning from multiple representations. *Learning and Instruction*, *13*, 227–237. http://doi.org/10.1016/S0959-47520200022-1

Shea, J. B., & Morgan, R. L. (1979). Contextual interference effects on the acquisition, retention, and transfer of a motor skill. *Journal of Experimental Psychology: Human Learning and Memory, 5*, 179–187. http://doi.org/10.1037//0278-7393.5.2.179

Slamecka, N. J., & Graf, P. (1978). The generation effect: Delineation of a phenomenon. *Journal of Experimental Psychology: Human Learning and Memory, 4*, 592–604. http://doi.org/10.3758/BF03198475

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. http://doi.org/10.1037/h0076934

Stalbovs, K., Scheiter, K., & Gerjets, P. (2015). Implementation intentions during multimedia learning: Using if-then plans to facilitate cognitive processing. *Learning and Instruction, 35*, 1–15. http://doi.org/10.1016/j.learninstruc.2014.09.002

Steele, B. (1991). Integrating art. *BCATA Journal for Art Teachers*, *31*, 41–44.

Stein, M., & Power, B. (1996). Putting art on the scientist’s palette. In R. S. Hubbard & K. Ernst (Eds.), *New entries: Learning by writing and drawing*. Portsmouth, NH: Heinemann.

Stieff, M., Hegarty, M., & Deslongchamps, G. (2011). Identifying representational competence with multi-representational displays. *Cognition and Instruction, 29*, 123–145. http://doi.org/10.1080/07370008.2010.507318

Sweller, J. (1999). *Instructional design in technical areas*. Camberwell, Australia: ACER Press.

Sweller, J. (2005a). Implications of cognitive load theory for multimedia learning. In R. E. Mayer (Ed.), *The Cambridge handbook of multimedia learning* (pp. 19–30). Cambridge, MA: Cambridge University Press.

Sweller, J. (2005b). The redundancy principle in multimedia learning. In R. E. Mayer (Ed.), *The Cambridge handbook of multimedia learning* (pp. 159–167). Cambridge, MA: Cambridge University Press.

Tirre, W. C., Manelis, L., & Leicht, K. L. (1979). The effects of imaginal and verbal strategies on prose comprehension by adults. *Journal of Reading Behavior*, *11*, 99–106.

van Essen, G., & Hamaker, C. (1990). Using student-generated drawings to solve arithmetic word problems. *Journal of Educational Research*, *83*, 301–312. http://doi.org/10.2307/27540404

van Meter, P. (2001). Drawing construction as a strategy for learning from text. *Journal of Educational Psychology*, *93*, 129–140. http://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. http://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, NC: 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. http://doi.org/10.1007/s10648-005-8136-3

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

Winne, P. H., & Hadwin, A. F. (1998). Studying as self-regulated learning. In D. J. Hacker, J. Dunlosky, & A. Graesser (Eds.), *Metacognition in educational theory and practice* (pp. 277–304). Hillsdale, NJ: Erlbaum.

Wittrock, M. C. (1990). Generative processes of comprehension. *Educational Psychologist*, *24*, 345–376. http://doi.org/10.1207/s15326985ep2404_2

Wittrock, M. C., & Alesandrini, K. (1990). Generation of summaries and analogies and analytic and holistic abilities. *American Educational Research Journal*, *27*, 489–502. http://doi.org/10.3102/00028312027003489

# 8 List of Figures

# 9 List of Tables